Archives of Gerontology and Geriatrics 94 (2021) 104337

Contents lists available at ScienceDirect

Archives of Gerontology and Geriatrics

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

Longitudinal impact and effects of booster sessions in a cognitive training program for healthy older adults

Lucas Matias Felix[a], Marcela Mansur-Alves[a], Mariana Teles[b,*], Laura Jamison[b], Hudson Golino[b]

[a] Universidade Federal de Minas Gerais, Department of Psychology: 6627 Antonio Carlos avenue, Belo Horizonte, Minas Gerais, 31270-901, Brazil
[b] University of Virginia, Deparment of Psychology: 485, McCormick Road, Gilmer Hall, Charlottesville, VA 22903, U.S.

ARTICLE INFO

Keywords:
Cognitive training
Long-term efficacy
Entropy
Aging

ABSTRACT

This paper reports the results from a 3-year follow-up study to measure the long-term efficacy of a cognitive training for healthy older adults and investigates the effects of booster sessions using an entropy-based metric. Design: semi-randomized quasi-experimental controlled design. Participants: 50 older adults, (M = 73.3, SD = 7.77) assigned into experimental (N = 25; Mean age = 73.9; SD = 8.62) and control groups (N = 25; mean age = 72.9; SD = 6.97). Instruments: six subtests of WAIS and two episodic memory tasks. Procedures: the participants were assessed on four occasions: after the end of the original intervention, pre-booster sessions (three years after the original intervention), immediately after the booster sessions and three months after the booster sessions. Results: the repeated measures ANOVA showed that two of the cognitive gains reported in the original intervention were also identified in the follow-up: Coding ($F(1, 44) = 11.79$, MSE = 0.77, $p = .001$, eta squared = 0.084) and Picture Completion ($F(1, 47) = 10.01$, MSE = 0.73, $p = .003$, eta squared = 0.060). After the booster sessions, all variables presented a significant interaction between group and time favorable to the experimental group (moderate to high effect sizes). To compare the level of cohesion of the cognitive variables between the groups, an entropy-based metric was used. The experimental group presented a lower level of cohesion on three of the four measurement occasions, suggesting a differential impact of the intervention with immediate and short-term effects, but without long-term effects.

1. Introduction

The area of cognitive training has accumulated considerable evidence towards the positive impact of training protocols on cognitive performance assessed immediately after the intervention (Martin, Clare, Altgassen, Cameron, & Zehnder, 2011; Nguyen, Murphy, & Andrews, 2019; Rebok et al., 2014). Most of the positive effects reported in the literature are in abilities that are part of the same cognitive domain trained during the intervention (trained skills) and for abilities that are not part of the cognitive domain trained, but that pertain to a closely related domain (near transfer effect).

However, the field has been challenged by a lack of evidence for long-lasting effects of cognitive gains after intervention. The most challenging aspect is keeping the cognitive gains for a long period of time (long-term efficacy) (Law, Barnett, Yau, & Gray, 2014; Nguyen et al., 2019). There has been a growing interest in investigating the long-term efficacy of cognitive training programs, especially due to the restricted number of long-term follow-up studies compared to the more common immediate effect studies. In a recent meta-analysis and systematic review, Nguyen et al. (2019) showed that, from a sample of 64 studies, only 16 examined the long-term efficacy of cognitive training programs, with the follow-up length ranging from 3 weeks to 18 months ($Mean_{months} = 6.83, SD = 5.09$).

One could argue that the lack of evidence for the long-term impact of cognitive interventions weakens the cognitive enrichment hypothesis (Salthouse, 2006) and that cognitive trainings have not been proven as capable to alter the rate of change in cognitive functioning throughout the lifespan (Salthouse, 2015). However, it's also possible that the lack of support for long-term effects of cognitive interventions may be attributable to methodological limitations and shortcomings in design or analysis that have been pointed to in the literature, such as the difficulty to control for multiple cofounders that can account for the age-related cognitive decline (Melby-Lervåg, Redick, & Hulme, 2016; Simons et al., 2016).

* Corresponding author at: 485 McCormick Road, Gilmer Hall, Room 102, Charlottesville, VA 22903, United States.
E-mail address: mt2yq@virginia.edu (M. Teles).

https://doi.org/10.1016/j.archger.2021.104337
Received 24 September 2020; Received in revised form 6 January 2021; Accepted 8 January 2021
Available online 13 January 2021

Alternatively, the exploration of the impact of an intervention should be extended beyond the analysis of the variability of cognitive scores before and after the training. This study presents a novel approach to investigate the long-term impact of a cognitive training in a group of healthy older adults with the use of an entropy-based metric termed *total correlation*. In doing so, patterns of change in the level of cohesion among the cognitive variables can be uncovered and provide relevant information for the impact of the training on the cognitive structure. This pattern of structural changes would be missed by traditional linear techniques that focus on the variability of the mean scores (i.e. whether or not they improve with and without the intervention).

In order to prolong the benefits of cognitive training, booster sessions may be conducted between posttest and follow-up testing occasions. The booster training is typically conceived as a shorter version of the original intervention. There is evidence supporting that additional training sessions throughout the follow-up period lead to a re-activation of the cognitive skills previously trained (e.g., Kelly et al., 2014; Nguyen et al., 2019) therefore being an important aspect to be considered when evaluating the long-term impact of the intervention.

The current study assessed a group of 50 older adults three years after a cognitive training had been conducted to investigate the long-term impact of this training. A booster training was included for the experimental group after the follow-up assessment, to explore the impact of additional training sessions in the cognitive performance. The level of cohesion among the cognitive variables was analyzed along the measurement occasions to investigate possible structural cognitive changes as a result of the intervention.

2. Literature review

Long-term efficacy is valuable evidence for the high quality and strength of a cognitive intervention program. Some of the most important and robust studies in the cognitive training field indicate that elderly people are able to retain the cognitive gains (training effects) from several months (Borella, Carretti, Riboldi, & De Beni, 2010; Brehmer, Westerberg, & Bäckman, 2012; Günther, Schäfer, Holzner, & Kemmler, 2003; Yang, Algesheimer, & Tessone, 2016) up to several years (Ball et al., 2002; Rebok et al., 2014; Willis et al., 2006).

The Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study shows strong long-term efficacy in improved performance of trained skills and the daily functioning of participants from a cognitive training program. According to the reported results (Rebok et al., 2014), the cognitive training gains for reasoning and processing speed were maintained for 10 years after the original intervention, and at least 5 years for memory, with a 4-session booster training at 11 and at 35 months after training.

The results of previous research suggest that the durability of cognitive gains for healthy elderly people involves multiple cognitive domains, such as memory, reasoning, processing speed and attention (Rebok et al., 2014; Yang et al., 2016). There is also support for the long-term efficacy of trainings focused on executive functions and daily functioning (Wilkinson & Yang, 2016).

When investigating the long-term efficacy of a cognitive training program, the inclusion of booster sessions, as well as the format and frequency of training sessions must be controlled. Booster sessions are typically conceived as a shorter version of a cognitive training program; their impact on the cognition can potentially lead to a re-activation of the cognitive skills previously trained. In meta-analysis studies conducted by Kelly et al. (2014) and Nguyen et al. (2019), the results showed that multidomain adaptive interventions that include booster sessions after the end of the training period can lead to an extension of the immediate gains and to a slower decline of the trained cognitive abilities. The authors argue that the additional tasks used in the booster sessions help maintain the cognitive strategies that were learned during the original training phase, strengthening the neuroplasticity mechanism in the aging process (Kelly et al., 2014; Nguyen et al., 2019).

However, it is important to observe that the inclusion of booster training sessions is not a condition for cognitive trainings to achieve long-term outcomes and benefits. There is evidence in the literature supporting the long-lasting impact of cognitive training programs even among studies that did not include booster sessions. A study conducted by Ball et al. (2002) found that the impact of cognitive training on the trained skills lasted for 2 years. Nguyen et al. (2019) also showed evidence of significant near and far-transfer training effects. The prolonged benefits of a cognitive intervention program are an important aspect to be investigated in the area of aging, especially given that the observed decline in cognitive functioning during the aging process negatively impacts peoples' autonomy and quality of life (Borella et al., 2010).

3. Research design

The current paper aims to assess a group of participants originally recruited and trained by Blinded for review # 1 three years after the first intervention. The goals are twofold. First, a traditional follow-up study after three years of the original intervention will compare the experimental and the control group to verify if the original gains of the former group reported by (Blinded for review #1) are maintained after a long interval of time. The second goal is to assess the effects of booster training by comparing mean improvement from baseline to post-booster training and three months after booster sessions between participants who did and did not receive booster training. A shorter version of the cognitive training program for elderly people developed by (Blinded for review #1) was conducted. This study follows a semi-randomized quasi-experimental controlled design. Considering previous studies assessing the durability effect of cognitive training on elderly people (e.g., Wilkinson & Yang, 2016) we hypothesize that the experimental group will keep surpassing the passive control group in terms of cognitive performance in all measurement occasions (pre-booster sessions, immediately after the booster sessions and the three-month follow-up).

3.1. Cognitive training

The original cognitive training consisted of twelve sessions of thirty minutes, conducted in an individual setting, once a week, using a series of paper-and-pencil tasks designed to train five domains: visual memory, perceptual-motor capacity, episodic memory, working memory and perceptual speed (for a detailed description of each session and its tasks see: Blinded for review #2).

The booster training used in the current study consisted of a reduced version of the original intervention, with eight sessions (twice a week, fifty minutes each). The tasks were selected based on their factor loading reported by (Blinded for review #3). Therefore, the booster training target four cognitive domains: perceptual speed, visual memory, episodic memory and working memory. A description of the booster training tasks can be seen in Table 1:

3.2. Sample and measures

The original sample consisted of 80 non-institutionalized individuals recruited from a Brazilian community (Blinded for review #1). The exclusion criteria adopted to recruit the sample were: age $\leq$ 60 years old; cognitive impairment detected through the Mini Mental State Examination. The normative study conducted by Brucki, Nitrini, Caramelli, Bertolucci, and Okamoto (2003) was used to detect individuals with cognitive impairment; depressive symptoms (Geriatric Depression Scale, score $>$ 7). A screening interview was used to exclude participants with a self-reported diagnosis of Alzheimer's disease, severe loss in vision and hearing or communicative ability. Participants with 30% or more absences in the training or cognitive assessment sessions were excluded from the final sample. The participants were randomly assigned to experimental or control groups. Some reallocations were made in order to balance age and education levels between the groups.

Table 1
Booster training: cognitive domain and task description.

Cognitive Domain	Task description
Perceptual speed	(1) Draw the correct path through a maze, without crossing over the lines, in a controlled time period; repeat the same maze in half the time required to complete it the first time; (2) Mark the stimulus-target within a series of distractor stimuli, with controlled time; repeat the task in half the time required for the first attempt.
Visual Memory	(1) Analyze figures and once they are gone, visualize them in your mind to answer questions about them; (2) Listen carefully to a story; split into segments; retell the story partially; retell the story in its entirety
Episodic Memory	(1) Apply different mnemonic strategies to memorize sequence of names; (2) Apply different mnemonic strategies to memorize sequence of numbers; (3) The instructor says a word and the participant should state a name that begins with the same syllable said by instructor
Working Memory	(1) Count the number of stimuli-targets in a set of distractor stimuli, while intoning a rhythm at the same time; (2) Read disorganized sections of a story to later retell it in the correct order, without help from the stimuli; (3) The instructor says a sequence of months and the participant should repeat it, ordering according to the calendar

This led to limitations on the randomization process but ensured that both groups had equivalent levels of age and educational level, two important factors that affect cognitive function.

From the original sample of 80 healthy individuals, 50 subjects were recruited for our follow up study and 30 participants were excluded for diverse drop-out reasons like death (n = 2), surgery recovery or domestic accident (n = 3), dementia diagnosis (n = 3), lack of interest to continue to participate (n = 3), moved to another city (n = 5), and unavailability to attend to training sessions (n = 17). The final sample was formed by 36 women and 14 men, ages 63 to 92 years old (M = 73.3; SD = 7.77; Median = 73). The recruitment procedure (from December 2017 to January 2018) used the same exclusion criteria adopted by (Blinded for review #1) and none of the 50 participants met the exclusion criteria. After the screening assessment they were allocated into the same experimental (N = 25; Mean age = 73.9 years; SD = 8.62) and control groups (N = 25; mean age = 72.9 years; SD = 6.97) as the original study (Golino, Mendoza, & Golino, 2017) . The experimental group has an average of 5.64 years of education (SD = 4.75), while the passive control group has an average of 5.88 years of education (SD = 4.36).

To check the exclusion criteria two instruments were used: (1) Geriatric Depression Scale (Sheikh & Yesavage, 1986, adapted by Almeida, 1999): participants with scores ≥ 7 were excluded from the sample; and (2) Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975), with participants scores ≥ 24 points being excluded (Valle, Castro-Costa, Firmo, Uchoa, & Lima-Costa, 2009). To assess the training effects, three instruments were used: (1) List recall task (Yassuda, Lasca, & Neri, 2005) contains a list of 35 grocery items, used to assess immediate recall.; (2) Story recall task (Yassuda et al., 2005) assesses the amount of ideas retrieved after reading a text; (3) Wechsler Adult Intelligence Scale (Wechsler, 1997) including six subtests: picture completion, coding; digit span, symbol search, arithmetic, and matrix reasoning.

3.3. Procedures

After proceeding with the participants screening, the exclusion criteria were checked and the participants (N = 50) were assigned to the experimental group (EG, N = 25) or the passive control group (CG, N = 25), in accordance with the original study (Blinded for review #1). For each measurement occasion (pre-booster sessions, immediately after the booster sessions and a three-month follow-up), the cognitive assessment consisted of one session per participant, that could be eventually extended for two sessions to accommodate the participant's availability and prevent the negative impact of tiredness on performance. The EG received eight booster sessions (60 minutes, once a week) and the CG participants did not receive any intervention. Both EG and CG were assessed with the same cognitive battery conducted along the same measurement occasions schedule: pre-booster sessions, immediately after booster sessions and three months after booster training.

3.4. Data analysis

To investigate the differences between groups (EG and CG), a repeated measures ANOVA was used, with group membership as the between subject effect and time as the within subject effect. The Shapiro-Wilk test was used to verify if the variables are normally distributed. Variables that did not present a normal distribution were transformed (i.e. normalized) using the bestNormalize package (Peterson & Cavanaugh, 2019). The effect size was calculated using generalized eta squared (see: Olejnik & Algina, 2003). All analysis were implemented in (R Core Team, 2017).

Additionally, an innovative approach based on entropy was used to measure the impact of the cognitive intervention. Entropy is a measure of uncertainty or disorder of a random variable (Shannon, 1948; Watanabe, 1960; Wiener, 1961). It can also be used to assess the degree of uncertainty (or disorder) of a set of variables, with lower values indicating lower uncertainty in the organization of the variables (Golino et al., 2020). Entropy measures were originally applied in physics (Watanabe, 1960), but its application has increased over the years in different areas, such as communication and neuroscience (Paninski, 2003), and psychology (Golino et al., 2020).

For a discrete random variable X, entropy is calculated using the distribution of $p(X)$ and is maximized when all events of X are equiprobable (i.e., $p(x) = \frac{1}{N}$, where N is the number of elements of X; Shannon, 1948). The most common method to estimate the entropy of a random variable is the maximum likelihood estimator (MLE) (Antos & Kontoyiannis, 2001). Let X be a discrete random variable that takes on values in a set $X = (x_1, x_2, \ldots, x_n)$, with a probability mass function $p(x) = P(X = x)$, the MLE estimator of entropy is:

$$H(X) = -\sum_{i=1}^{n} p(x) \log p(x)$$

Entropy was popularized after the publication of the mathematical theory of communication by Shannon (1948). Years earlier, however, Watanabe (1939) developed a modified index of entropy termed *total correlation* to measure not only the uncertainty of random variables, but also the strength of their correlation beyond their average interaction (Watanabe, 1960). The total correlation of a set of variables $\mathscr{X} = (x_1, x_2, \ldots, x_n)$ is calculated as follows (Watanabe, 1960):

$$C_{tot_{\mathscr{X}}} = \left(\sum_{i=1}^{n} H(x_i)\right) - H(x_1, x_2, \ldots, x_n) \geq 0$$

Considering two sets of variables, υ and ω, the difference between their joint entropy and the sum of their individual entropies is a measure of the correlation between them (Watanabe, 1960). As pointed by Golino et al. (2020), total correlation captures the decrease of ignorance of ω provided by the observation of υ, being a non-linear measure of uncertainty and association.

An example illustrates well what total correlation captures, and why using it to investigate the impact of cognitive interventions can provide useful information that traditional linear and parametric statistical techniques can't. Suppose we have four random normal variables (A, B, C and D). In a baseline condition these variables present the same mean ($M = 15$) and standard deviation ($\sigma = 5$), and their correlation is:

$$Cor = \begin{bmatrix} 1 & 0.4 & 0.5 & 0.6 \\ 0.4 & 1 & 0.4 & 0.5 \\ 0.5 & 0.4 & 1 & 0.3 \\ 0.6 & 0.5 & 0.3 & 1 \end{bmatrix}$$

Now, this baseline condition can be compared to four different conditions, varying the means and the correlation between A, B, C and D. In the first condition (*high correlation*), the means and variances of the variables are the same as the baseline condition, but variable A presents a higher correlation with the remaining variables (B, C and D): 0.6, 0.7 and 0.8, respectivelly. The second condition (*low correlation*) is the same as the first, but the correlation between A and B, C and D is now set to 0.1, 0.2, 0.3, respectively. In the third condition (*high mean*), the four variables present the same correlation and variances as the baseline condition, but now variable A has a mean of 25. Finally, the fourth condition (*low mean*) is similar to the third, but the mean of A is now set to 5.

To illustrate what happens with the total correlation in each condition (including the baseline), a brief simulation was implemented. For each condition, 1000 datasets with four variables (A, B, C and D) and 1000 observations were generated. The data generation mechanism used is described next. For each simulated dataset, four independent random normal variables with mean μ and standard deviation σ were generated with $N = 1000$, generating a data matrix D. The simulated dataset has variables that are independent from each other, therefore the correlation among the variables is zero. To obtain a transformation of the variables with the same mean (μ) and standard deviation σ, but with a desired correlation matrix **Cor**, a simple transformation can be implemented:

$$D_{New} = D \times \mathbf{Cor}_{Cholesky} + \sigma + \mu$$

Where D is the original data (with four independent normal random variables and 1000 observations), $\mathbf{Cor}_{Cholesky}$ is the Cholesky decomposition of the desired correlation matrix, μ and σ are the mean and standard deviation of the original data, and D_{New} is the new data with 1000 observations and four normal random variables with correlation matrix **Cor**.

Fig. 1 presents the mean and the distribution of the total correlation calculated in each simulated dataset, per condition (left side) as well as the 95% confidence interval of the total correlation means (right side). As expected, total correlation changes with the change in the correlation between variables and with the change in their means. The total correlation of the variables increases if the correlation increases, but the mean and standard deviation is kept the same. On the other side, total correlation decreases if the correlation is lower and the mean and standard deviation remains unchanged. When the correlation is the same, but the means increases or decreases, total correlation decreases.

An analysis of variance would indicate a significant difference between the baseline conditions and the high and low mean conditions, only. But would suggest that conditions one and two (high and low correlation) are not statistically different from the baseline. Compared to a traditional analysis of variance, total correlation offers at least two important advantages according to (Golino et al., 2020); firstly, it is a non-parametric and non-linear metric, not requiring the data to be normally distributed and enabling the measurement of non-linear relations. Secondly, it varies not only with the variation of the mean, but also with the variation of the correlation between variables. This is an important characteristic that can give additional relevant information to studies investigating the impact of cognitive interventions. How the level of cohesion among the cognitive abilities vary as a result of the intervention may add significant information about the impact of the training on the cognitive structure as a whole. This therefore allows patterns that would not be possible to be detected only with the analysis of the variability of the cognitive scores to be uncovered.

Total correlation was estimated using the EGAnet package (Golino & Christensen, 2019). To calculate the 95% confidence interval of total correlation, a boostrap approach (Chihara & Hesterberg, 2011) was used in which the observed data is resampled without replacement 1000 times with 20 observations each. The codes used in the current paper, the data and the RMarkdown file used to combine text, data and code is available in an Open Science Framework repository: https://osf.io/b35eq/?view_only=812f21139c30411095d8f0e3fea14dd7. The current research received approval from the local IRB.

4. Ethical standards

The authors assert that all procedures contributing to this work

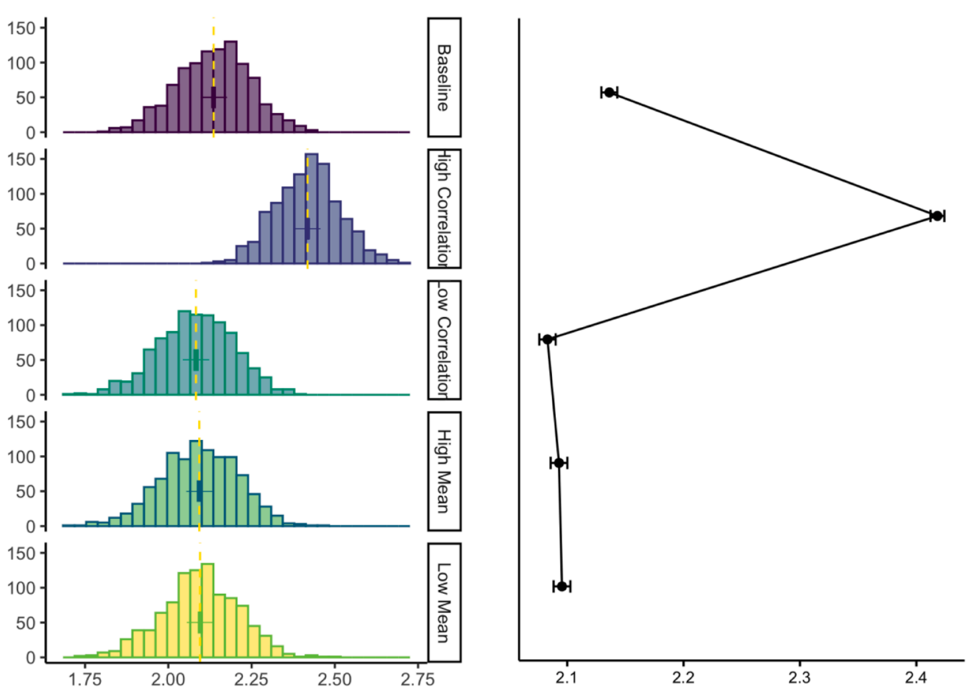

Fig. 1. Total correlation per condition tested.

comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

5. Results

The demographic characteristics of the sample are depicted in Table 2. Mean and standard deviation were calculated for the continuous variables and the categorical variables are expressed in percentage. None of the participants received cognitive training during the interval between the original intervention and the follow-up assessment. Both groups had similar baseline cognitive performance (MMSE) and no significant differences were found for level of education, daily life functioning, social economic status, and marital status.

The repeated measures ANOVA showed that three years after the original intervention, but before the booster training, the participant's scores were maintained for two of the three gains reported by Golino et al. (2017). There was a significant interaction for time and group membership for coding ($F(1,44) = 11.79$, $MSE = 0.77$, $p = 0.001$, $\hat{\eta}_G^2 = 0.084$) and picture completion ($F(1, 47) = 10.01$, $MSE = 0.73$, $p = 0.003$, $\hat{\eta}_G^2 = 0.060$), with moderate effect sizes. Notably, a significant interaction was also found for symbol search ($F(1,38) = 7.12$, $MSE = 0.77$, $p = 0.011$, $\hat{\eta}_G^2 = 0.055$), arithmetic ($F(1,47) = 5.21$, $MSE = 0.61$, $p = 0.027$, $\hat{\eta}_G^2 = 0.030$), story recall task ($F(1,37) = 4.64$, $MSE = 0.58$, $p = 0.038$, $\hat{\eta}_G^2 = 0.044$) and memory list ($F(1,37) = 7.29$, $MSE = 0.70$, $p = 0.010$, $\hat{\eta}_G^2 = 0.084$), which was not found in the original study. No considerable interaction effect was found for digit span ($F(1,47) = 2.57$, $MSE = 0.80$, $p = 0.116$, $\hat{\eta}_G^2 = 0.022$) and matrix reasoning ($F(1,47) = 2.74$, $MSE = 0.75$, $p = 0.105$, $\hat{\eta}_G^2 = 0.019$).

However, it is important to point out that the time/group effect found does not favor the experimental group (Fig. 2). After three years from the original intervention and before booster training, both groups presented a decline in their average performance, but the experimental group had the sharpest decline. Surprisingly, the control group surpassed the experimental group in most of the cognitive variables measured at this occasion. The scenario changes after the booster sessions. While the control group continued in its trajectory of decline, the experimental group showed improved performance in all variables measured.

Table 2
Characteristics of the sample (Follow-up of 3 years).

Variable	*n*	%	Groups EG (25)	CG (25)	W[b] (Sig.)
Age	50		73,9 (8,62)	72,9 (6,97)	329 (0,74)
Years of education	50		5,64 (4,75)	5,88 (4,36)	290 (0,66)
MMSE	50		25,9 (4,16)	25,6 (3,50)	318 (0,92)
GDS	50		3,36 (1,80)	2,96 (1,71)	346 (0,51)
IADL	50		14,2 (6,36)	12,7 (5,07)	353 (0,43)
Number of children	50		5,12 (3,63)	3,64 (2,30)	241 (0,16)
Social Economic Status[c]					
Class B	8	16%	4 (16%)	4 (16%)	
Class C	31	62%	18 (72%)	13 (52%)	329 (0,74)
Class D	11	22%	3 (12%)	8 (32%)	
Marital Status					
Single	4	8%	2 (8%)	2 (8%)	
Married	19	38%	10 (40%)	9 (36%)	276 (0,45)
Divorced / Separated	4	8%	3 (12%)	1 (4%)	
widow	21	42%	10 (40%)	11 (44)	
Others	2	4%		2 (8%)	
Have you received any non-pharmacological intervention in the last 3 years?[a]					
Not	46	92%	24 (96%)	22 (88%)	287 (0,30)
Yes (psychotherapy)	4	8 %	1 (4%)	3 (12%)	

Note. MMSE = Mini Mental State Examination; GDS = Geriatric Depression Scale; IADL = Instrumental Activities of Daily Living Scale (national adaptation by Santos & Virtuoso-Junior, 2008); EG = experimental group; CG = control group.

[a] Question included in the follow-up study interview.
[b] Test Wilcoxon-Mann-Whitney
[c] According to Brazilian criterion for economic classification (Brasil, 2008)

The results point to a positive short-term effect of the booster sessions (Table 3), since all variables presented a significant interaction between group and time, with a large effect sizes for symbol search ($F(1, 47) = 23.40$, $MSE = 0.48$, $p < 0.001$, $\hat{\eta}_G^2 = 0.167$) and story recall task ($F(1,47) = 34.56$, $MSE = 0.16$, $p < 0.001$, $\hat{\eta}_G^2 = 0.148$).

The results of the follow-up assessment three months after the booster session (Table 4) indicate that the differences found between the experimental and the control group were not just maintained (since all variables presented a significant interaction between group and time) but actually increased. The effect sizes ranged from moderate to high, as shown in Table 4: Matrix reasoning ($F(1, 47) = 24.18$, $MSE = 0.21$, $p < 0.001$, $\hat{\eta}_G^2 = 0.073$), arithmetic ($F(1, 47) = 33.07$, $MSE = 0.14$, $p < 0.001$, $\hat{\eta}_G^2 = 0.093$), digit span ($F(1, 47) = 21.27$, $MSE = 0.24$, $p < 0.001$, $\hat{\eta}_G^2 = 0.103$) presented moderate effect sizes, while picture completion ($F(1, 47) = 30.47$, $MSE = 0.43$, $p < 0.001$, $\hat{\eta}_G^2 = 0.164$), coding ($F(1,47) = 33.01$, $MSE = 0.29$, $p < 0.001$, $\hat{\eta}_G^2 = 0.186$), symbol search ($F(1,47) = 24.59$, $MSE = 0.57$, $p < 0.001$, $\hat{\eta}_G^2 = 0.207$), memory list ($F(1, 47) = 47.68$, $MSE = 0.27$, $p < 0.001$, $\hat{\eta}_G^2 = 0.247$) and story recall task ($F(1, 47) = 66.14$, $MSE = 0.22$, $p < 0.001$, $\hat{\eta}_G^2 = 0.342$) presented large effect sizes.

Fig. 2 shows the mean and 95% confidence interval of the scores on each cognitive variable, per group and measurement occasion.

5.1. Total correlation per group

Fig. 3 shows the mean and 95% confidence interval of total correlation per group and measurement occasion (including the post-test of the original study). The experimental group presented a lower total correlation in comparison to the control group in the post-test of the original study, in the post-test of the booster sessions and in the 3-month follow up. This indicates a lower level of cohesion among the cognitive measures for the experimental group.

6. Discussion

This study assessed a subsample of 50 participants three years after the original intervention they were a part of conducted by Golino et al. (2017). The original study reported significant differences between EG and CG (time/group interaction effects) for three variables. After three years but before booster training, significant differences were maintained for two of them (coding and picture completion), plus four other variables (symbol search, arithmetic, list recall task, and story recall task).

However, the time/group effect found favored the control group, that surpassed the experimental group in most of the cognitive variables measured in the three-year follow-up. The experimental group presented the strongest decline in terms of cognitive performance. This result shows that the original intervention did not present a long-term impact, contrasting with classical studies that show a longer-lasting cognitive gains (training effects) ranging from months (Borella et al., 2010; Brehmer et al., 2012; Yang & Krampe, 2009) to several years (Ball et al., 2002; Rebok et al., 2014; Willis et al., 2006) The pattern reverses after the booster sessions, where the experimental group showed improved performance in all cognitive measures assessed immediately after the booster sessions compared with the passive control group.

Most of the literature reports a loss of the cognitive gains after the end of the cognitive intervention programs (Hertzog, Kramer, Wilson, & Lindenberger, 2008; Kounti et al., 2011; Nguyen et al., 2019). Hertzog

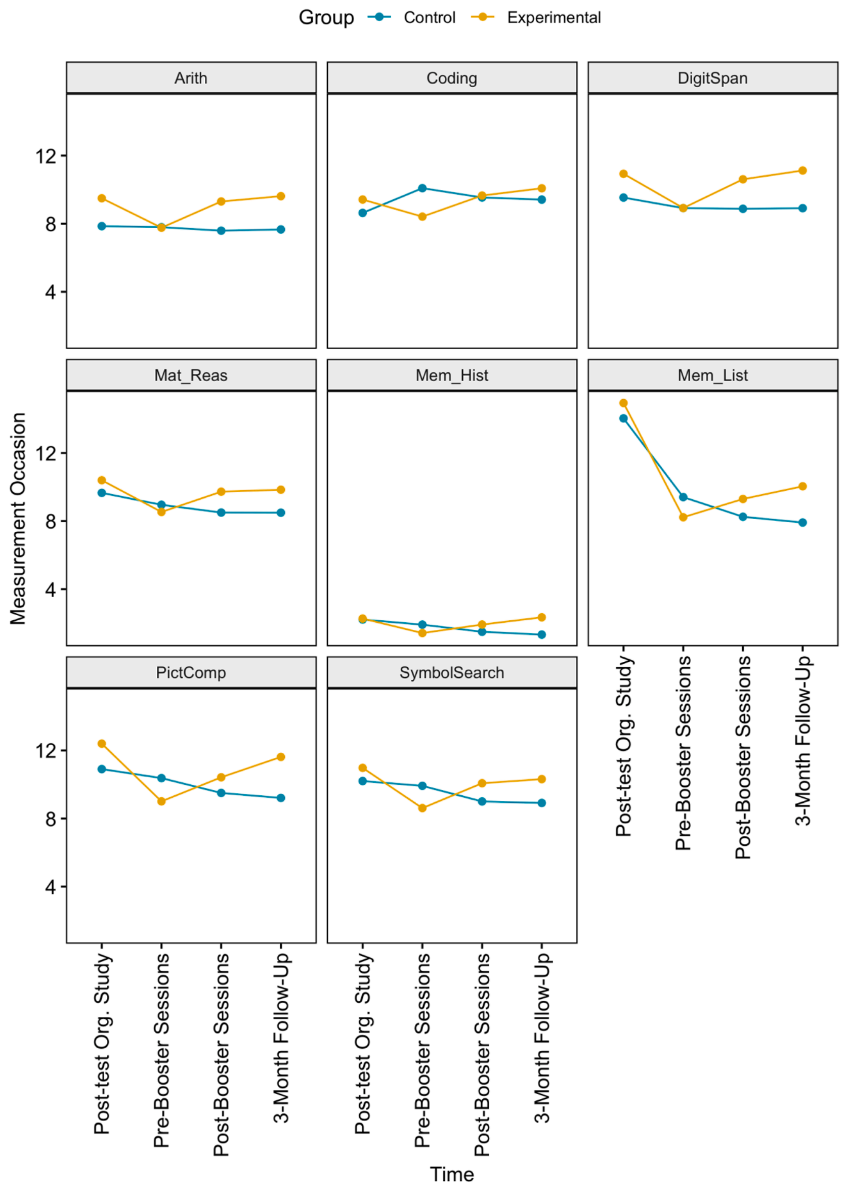

Fig. 2. Resampled means and confidence intervals for each variable, in each measurement point.

et al. (2008) argue that once the set of stimuli that were supporting the cognitive responses is taken away, the performance levels return to the baseline performance, or close to it. The results of the present study are in accordance with this expected pattern, with no long-lasting impact demonstrated in the absence of booster sessions. Similar results have been reported by the literature [Ball et al. (2002); Martin et al. (2011); Rebok et al. (2014); willis et al., 2006long], supporting the argument that the durability effects are stronger with (or even conditional on) the inclusion of booster sessions.

In cognitive training studies, it is expected that some skills or abilities will improve due to the intervention. However, not all skills or abilities trained will improve. If we quantify only the average variation per skill or ability (or construct measured pre- and post-intervention), we might be losing information regarding the level of cohesion of the variables. A higher cohesion means that the distribution of the variables (or scores) per person is more uniform. For example, if people are improving all their skills or abilities, then they will present higher scores for every measured variable.

Conversely, if people are losing their skills or abilities, then they will present lower scores in all measured variables. In both cases, the cohesion of the variables is high. However, some skills or abilities may be improving, while other remain unchanged. In this case, the level of cohesion among the variables measured will be lower, because the distribution of the scores per individual will not be very uniform. Therefore, in cognitive training studies, we can expect that the cohesion among variables measured pre and post-intervention should decrease for the experimental group (compared to the control group), if only a subset of skills or abilities are being impacted by the training protocol. Total correlation, as mentioned earlier, is a non-linear metric that can capture the level of cohesion among variables.

Therefore, it is expected that the gains in performance will negatively impact the cohesion of the cognitive variables, since the cognitive training can have an impact on two different fronts. A successful cognitive intervention will generate an increase in performance on cognitive measures, leading to higher mean scores for the experimental group in comparison with the control group. At the same time, the correlation between variables may be lower for the experimental group, since the gains in performance are not equal for every cognitive measure used. As a consequence, the experimental group may present lower total correlations, or lower cohesion.

Table 3
ANOVA - Immediate short-term effects of the booster session (pre vs. post-booster sessions).

Variables	Effect	*F*	df_1^{GG}	df_2^{GG}	*MSE*	*p*	$\hat{\eta}^2_G$
Digit Span	Group	8.17	1	47	0.63	.006	.121
	Time	27.03	1	47	0.16	< 0.001	.105
	Group x Time	18.92	1	47	0.16	< 0.001	.076
Coding	Group	5.44	1	47	0.74	.024	.079
	Time	3.61	1	47	0.26	.064	.020
	Group x Time	19.28	1	47	0.26	< 0.001	.097
Picture Completion	Group	0.35	1	47	1.31	.556	.006
	Time	4.91	1	47	0.28	.032	.018
	Group x Time	20.11	1	47	0.28	< 0.001	.069
Symbol Search	Group	0.04	1	47	0.71	.850	.000
	Time	2.48	1	47	0.48	.122	.021
	Group x Time	23.40	1	47	0.48	< 0.001	.167
Arithmetic	Group	3.38	1	47	0.93	.073	.058
	Time	12.65	1	47	0.16	.001	.037
	Group x Time	20.93	1	47	0.16	< 0.001	.060
Matrix Reasoning	Group	1.62	1	47	1.14	.209	.028
	Time	5.03	1	47	0.21	.030	.016
	Group x Time	21.77	1	47	0.21	< 0.001	.067
Memory - List	Group	0.11	1	47	0.81	.741	.002
	Time	0.44	1	47	0.15	.510	.001
	Group x Time	44.33	1	47	0.15	< 0.001	.129
Memory - History	Group	0.01	1	47	0.53	.914	.000
	Time	0.29	1	47	0.16	.590	.001
	Group x Time	34.56	1	47	0.16	< 0.001	.148

Table 4
ANOVA - three month follow-up of the booster session (pre-booster sessions vs. three-month follow-up).

Variables	Effect	*F*	df_1^{GG}	df_2^{GG}	*MSE*	*p*	$\hat{\eta}^2_G$
Digit Span	Group	10.96	1	47	0.70	.002	.148
	Time	29.66	1	47	0.24	< 0.001	.138
	Group x Time	21.27	1	47	0.24	< 0.001	.103
Coding	Group	2.33	1	47	0.60	.134	.032
	Time	6.18	1	47	0.29	.017	.041
	Group x Time	33.01	1	47	0.29	< 0.001	.186
Picture Completion	Group	0.32	1	47	0.98	.574	.005
	Time	9.94	1	47	0.43	.003	.060
	Group x Time	30.47	1	47	0.43	< 0.001	.164
Symbol Search	Group	0.10	1	47	0.57	.755	.001
	Time	2.81	1	47	0.57	.100	.029
	Group x Time	24.59	1	47	0.57	< 0.001	.207
Arithmetic	Group	5.48	1	47	0.80	.023	.091
	Time	23.67	1	47	0.14	< 0.001	.069
	Group x Time	33.07	1	47	0.14	< 0.001	.093
Matrix Reasoning	Group	1.87	1	47	1.16	.177	.033
	Time	6.89	1	47	0.21	.012	.022
	Group x Time	24.18	1	47	0.21	< 0.001	.073
Memory - List	Group	0.91	1	47	0.57	.344	.013
	Time	1.94	1	47	0.27	.170	.013
	Group x Time	47.68	1	47	0.27	< 0.001	.247
Memory - History	Group	5.09	1	47	0.38	.029	.064
	Time	3.37	1	47	0.22	.073	.026
	Group x Time	66.14	1	47	0.22	< 0.001	.342

It is well established in the literature that a differential pattern of change among the cognitive abilities is expected throughout the life-span. Some variables decline over time with different onsets of performance decrease, while others may stabilize or even improve over time, as in the case of vocabulary (Salthouse, 2019). Therefore, it is expected that the level of cohesion among the cognitive variables (or their total correlation) can decline with time, since the performance in some variables decline, while others remain stable or improve with time.

In our study, the three-year interval revealed that the control group presented a decrease in the level of cohesion, possibly due to the expected differential aging effects on the cognitive variables used here. The trajectory of the means for the control group showed that some of the variables declined over time as others improved, which could possibly explain the loss of cohesion among the variables. The experimental group, on the other hand, had a slight increase in cohesion during the three-year interval, possibly because the gains in performance due to the original intervention faded, so the performance in the cognitive variables returned to its original levels (i.e., declined during this interval). It's possible to observe that all variables showed a similar pattern of change in performance (means) during this three-year interval, leading to a higher cohesion in the cognitive structure for the experimental group.

The pattern of the cohesion levels was investigated using two strategies. Combined, the mean difference between the highest and the lowest score for each participant, and the difference between the highest and the lowest average correlation points to a direction suggested by the brief simulation shown on Fig. 1. In sum, cohesion decreases both if there is variability in the means of the variables (decreasing or increasing) and if the linear correlation between variables decreases. That is exactly what happened with the experimental group. In three different measurement occasions, i.e., the post-test of the original study and in the two post-tests of the booster sessions (post-booster sessions and 3-month follow-up), the experimental group presented higher means then the control group, but also lower correlation between variables. The higher means and lower correlations, captured well by the non-linear index of total correlation, are complementary supporting evidence for the idea that the booster sessions impacted the cognitive structure of the experimental group.

Computing the difference between the highest and the lowest score for each participant in each measurement occasion can shed some light into why the experimental group presents a lower level of cohesion than the control group. Fig. 4 shows that experimental group presents a higher mean difference between the highest and the lowest score for each participant in comparison to the control group in the post-test of the original study and in the two post-tests of the booster sessions (post-booster sessions and three-month follow-up). This is due to a greater level of dispersion of the scores for the experimental group, which leads to a decreased level of cohesion compared to the control group.

At the same time, computing the difference between the maximum and the minimum average correlation between variables for each group, in each measurement point, can also help us understand the results of total correlation. Fig. 5 shows that the experimental group has a lower mean difference of average correlation in the post-test of the original study and in the two post-tests of the booster sessions. This indicates that the cognitive variables in the experimental group are less strongly related and that the difference between the variable with the highest and the lowest average correlation is much smaller compared to the control group. The mean and 95% confidence intervals depicted in Fig. 5 were computed using a bootstrap approach (Chihara & Hesterberg, 2011) in

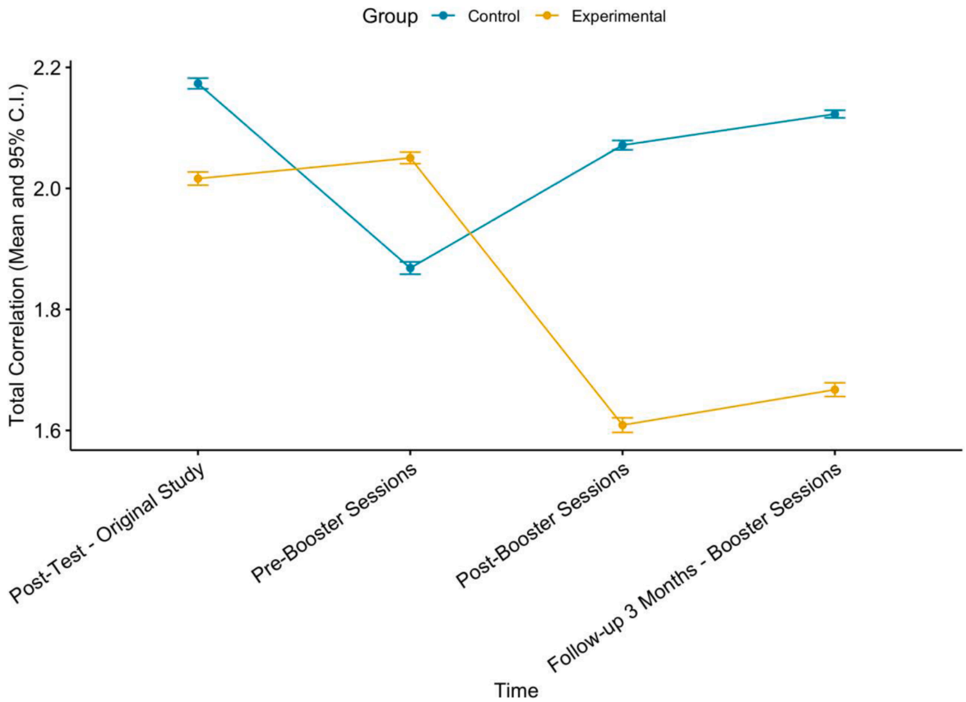

Fig. 3. Total correlation per group.

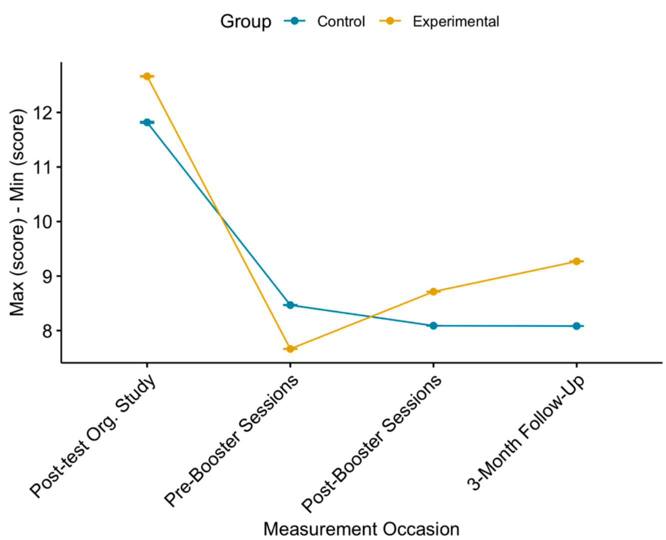

Fig. 4. Average difference (resampled means and confidence interval) between the maximum and the minimum score across variables per group in each measurement point.

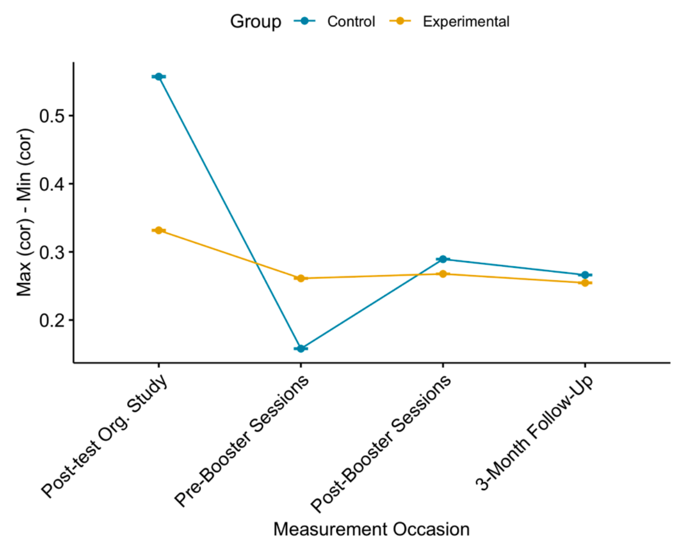

Fig. 5. Average difference (resampled means and confidence interval) between the maximum and the minimum average correlation per group in each measurement point.

which the observed data is resampled without replacement 1000 times with 20 observations each.

At the same time, the level of cohesion increased for the control group in the two post-booster sessions measurements. At first sight this may seem a strange pattern, since the control group continued in its downward trajectory (in terms of the means of the variables used in this study). But the simulation presented briefly in this paper explains this pattern. The level of cohesion among variables (i.e., total correlation), increases if one or more variables presents a higher correlation with the other variables, leading to an increase in the difference between the highest and the lowest average correlation. The difference in the average correlation shown in Figure is exactly what happens with the control group.

Important limitations of our study are the small sample size and the use of a passive control group, conditions that limit our capacity to account for test-retest/practice effects (Simons et al., 2016). Small sample sizes also reduce the likelihood that a significant result reflects a true effect in population (Button et al., 2013). Given the relatively small sample size of our study, it is possible that differences between EG and CG might have influenced the reported effects. The difference in the frequency of sessions between the original and booster trainings - once a week in the original and twice a week in the booster training – is another limitation of our design that may impact the interpretation of the results as a greater efficacy of EG post-booster training may represent a function of training volume. More volume-matched comparative studies are required in the field of cognitive training to explore this question.

An important implication of this study is to introduce a novel approach to investigate how a cognitive training can impact the level of cohesion among the cognitive variables throughout the measurement occasions. We used total correlation as an entropy metric to detect how the cognitive variables are connected among them and how the pattern of structural changes observed between the experimental and control groups can be interpreted and attributable to the intervention. The field of cognitive training typically rely on linear and multiple analysis of variance and covariance to investigate the differences in improvement between the experimental and control groups (Simons et al., 2016). This analytical approach, however, does not detect possible structural changes in the cohesion of the trained skills that would be attributable to a cognitive training. The authors agree that this is a non-neglectable information that should be considered by the field as an important training effect to be investigated.

7. Conclusion

In the current paper we show that not only the booster sessions improved people's cognitive performance, on average, but it reversed the trajectory of cognitive decline due to a regression to the baseline cognitive performance. Our result endorses the growing agreement in the field that the long-term efficacy of a training is reachable when associated with booster sessions. The entropy-based metric total correlation was used as a novel approach to analyze the long-term effects of cognitive training on the level of cohesion of cognitive structure.

Conflict of Interest declaration and funding statement

The authors wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

CRediT authorship contribution statement

Lucas Matias Felix: Methodology, Investigation. **Marcela Mansur-Alves:** Conceptualization, Methodology, Supervision. **Mariana Teles:** Writing - original draft. **Laura Jamison:** Writing - review & editing. **Hudson Golino:** Formal analysis.

Declaration of Competing Interest

None.

References

Almeida, O. P., & Almeida, S. A. (1999). Confiabilidade da versão brasileira da escala de depressão em geriatria (gds) versão reduzida [*Reliability study of Brazilian version of the geriatric depression scale, reduced version*]. *Arquivos de Neuro-Psiquiatria, 57*(2B), 421–426.

Antos, A., & Kontoyiannis, I. (2001). Convergence properties of functional estimates for discrete distributions. *Random Structures & Algorithms, 19*(3–4), 163–193. https://doi.org/10.1002/rsa.10019

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., & Marsiske, M. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA, 288*(18), 2271–2281.

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging, 25*(4), 767–778.

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 6*(63). https://doi.org/10.3389/fnhum.2012.00063

Brucki, S., Nitrini, R., Caramelli, P., Bertolucci, P. H., & Okamoto, I. H. (2003). Sugestoes para o uso do mini-exame do estado mental no brasil. *Arquivos de Neuro-Psiquiatria, 61*(3B), 777–781.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365–376.

Chihara, L., & Hesterberg, T. (2011). *Mathematical statistics with resampling and R*. Wiley Online Library.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Golino, M. T. S., Mendoza, C. F., & Golino, H. F. (2017). Effects of cognitive training on cognitive performance of healthy older adults. *The Spanish Journal of Psychology, 20*(39). https://doi.org/10.1017/sjp.2017.38

Golino, H., & Christensen, A.P. (2019). EGAnet: Exploratory graph analysis: A framework for estimating the number of dimensions in multivariate data using network psychometrics. Retrieved from https://CRAN.R-project.org/package=EGAnet.

Golino, H., Moulder, R., Shi, D., Christensen, A., Garrido, L., Neto, M., & Boker, S. (2020). Entropy fit indices: New fit measures for assessing the structure and dimensionality of multiple latent variables. *Multivariate Behavioral Research*. https://doi.org/10.1080/00273171.2020.1779642

Günther, V. K., Schäfer, P., Holzner, B., & Kemmler, G. (2003). Long-term improvements in cognitive performance through computer-assisted cognitive training: A pilot study in a residential home for older people. *Aging & Mental Health, 7*(3), 200–206.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest, 9*(1), 1–65.

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

Kounti, F., Bakoglidou, E., Agogiatou, C., Lombardo, N. B. E., Serper, L. L., & Tsolaki, M (2011). RHEA,* a nonpharmacological cognitive training intervention in patients with mild cognitive impairment: A pilot study. *Topics in Geriatric Rehabilitation, 27*(4), 289–300. https://doi.org/10.1097/TGR.0b013e31821e59a9. https://doi.

Law, L. L., Barnett, F., Yau, M. K., & Gray, M. A. (2014). Effects of functional tasks exercise on older adults with cognitive impairment at risk of alzheimer's disease: A randomised controlled trial. *Age and Ageing, 43*(6), 813–820.

Martin, M., Clare, L., Altgassen, A. M., Cameron, M. H., & Zehnder, F. (2011). *Cognition-based interventions for healthy older people and people with mild cognitive impairment*. The Cochrane Library.

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.

Nguyen, L., Murphy, K., & Andrews, G. (2019). Immediate and long-term efficacy of executive functions cognitive training in older adults: A systematic review and meta-analysis. *Psychological Bulletin, 145*(7), 698–733.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*(4), 434.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Computation, 15*(6), 1191–1253. https://doi.org/10.1162/089976603321780272

Peterson, R. A., & Cavanaugh, J. E. (2019). Ordered quantile normalization: A semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 1–16. https://doi.org/10.1080/02664763.2019.1630372

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., & King, J. W. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society, 62*(1), 16–24.

Salthouse, T. A. (2006). Mental exercise and mental aging: Evaluating the validity of the "use it or lose it" hypothesis. *Perspectives on Psychological Science, 1*(1), 68–87.

Salthouse, T. A. (2015). Do cognitive interventions alter the rate of age-related cognitive change? *Intelligence, 53*, 86–91.

Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging, 34*(1), 17.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(3), 379–423.

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric depression scale (gds): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17*(3), 103–186.

Valle, E. A., Castro-Costa, E., Firmo, J. O., Uchoa, E., & Lima-Costa, M. F. (2009). A population-based study on factors associated with performance on the mini-mental state examination in the elderly: The bambui study. *Cadernos de Saude Publica, 25*(4), 918–926.

Watanabe, S. (1939). Über die anwendung thermodynamischer begriffe auf den normalzustand des atomkerns. *Zeitschrift Für Physik, 113*(7-8), 482–513.

Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development, 4*(1), 66–82. https://doi.org/10.1147/rd.41.0066

Wechsler, D. (1997). *Wechsler adult intelligence scale - wais-iii*. Psychological Corporation.

Wiener, N. (1961). *Cybernetics or control and communication in the animal and the machine, 25*. MIT Press.

Wilkinson, A. J., & Yang, L. (2016). Long-term maintenance of inhibition training effects in older adults: 1-and 3-year follow-up. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 71*(4), 622–629.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., & Koepke, K. M. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA, 296*(23), 2805–2814.

Yang, L., & Krampe, R. T. (2009). Long-term maintenance of retest learning in young old and oldest old adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 64*(5), 608–611.

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports, 6*, 30750.

Yassuda, M. S., Lasca, V. B., & Neri, A. L. (2005). Meta-memoria e auto-eficacia: Um estudo de validação de instrumentos de pesquisa sobre memoria e envelhecimento. *Psicologia: Reflexão E Critica, 18*(1), 78–90.