PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The effect of Computerized Cognitive Training on cognitive outcomes in Mild Cognitive Impairment: A Systematic Review and Meta-Analysis
AUTHORS	Zhang, Haifeng; Huntley, Jonathan; Bhome, Rohan; Holmes, Benjamin; Cahill, Jack; Gould, Rebecca; Wang, Huali; Yu, Xin; Howard, R

VERSION 1 - REVIEW

REVIEWER	Isabel Elaine Allen University of California San Francisco
REVIEW RETURNED	28-Oct-2018

GENERAL COMMENTS	Overall a very comprehensive & well written systematic review. A couple of suggestions: 1. The meta-analysis needs to be registered on Prospero to ensure that the protocol & results are available to other researchers for replication and updating. It is the only thing missing from the Prisma checklist. 2. The authors should perform their search on PubMed to ensure that no studies were missed as their are different subject matter articles on PubMed compared to Web of Science. 3. Table 1 is excellent but too comprehensive to be included in the main body of the paper - perhaps shorten the summary for each study and move the majority of the table to an online supplement.

REVIEWER	Jeremy Silverman Department of Psychiatry Icahn School of Medicine at Mount Sinai New York, NY, USA
REVIEW RETURNED	14-Dec-2018

GENERAL COMMENTS	The paper is a systematic review and meta-analysis of studied assessing the efficacy of computerized cognitive training for late middle aged and elderly people with mild cognitive impairment. The paper is well-written and carefully conducted, and their methods and analysis are effectively described with good and appropriate detail. However, given the recent review by Hill, Mowszowski, Naismith et al. (Computerized cognitive training in older adults with mild

	cognitive impairments or Dementia: a systematic review and meta-analysis. American Journal of Psychiatry, 2017; 174:329-340), a question that needs to be more fully addressed is what is added contribution this paper provides. The authors cite the earlier review and other related ones, but they need to show what their report adds to the question of CCT efficacy beyond what has already been published. Other very minor concerns are itemized below. 1. In the Abstract and the Introduction, the authors write that their objective is to "determine the efficacy" (line 37) and that it "investigates the efficacy" (line 133) of CCT for MCI, but as this is not a direct study of CCT, but a review of other studies, this language might be revised. 2. Line 124-5 states "Critical analysis of research using for MCI should reveal insight into any effective components of CCT…." This was somewhat unclear. Perhaps what is meant is, "Ideally, critical analysis…would reveal insight into specific components…"? 3. Line 245 discusses the effect sizes found in studies of global cognition and each of the cognitive domains using active control groups versus those with non-active controls. While the stats and significance of the CCT versus control effect sizes are provided, it would also be useful to test the differences between the effect sizes. For example, beginning on line 254, it is noted that for studies of memory using active controls, the "effect size was statistically significant and larger than that of trials with passive control groups" (the latter was not significant). Might these different effect sizes be tested for statistical significance? 4. In another subgroup analysis – multi-domain versus single domain CCT – the g statistic was significant for multi-domain CCT, but not for single domain CCT. However, the g statistic – was nominally larger for the single domain CCT (.31) than the multi-domain CCT (.30). This suggests that the difference in significance is probably a function of the reduced power in the single domain studies. A test comparing these two g-statistics would surely be non-significant and would show that, at least based on the available studies to date, there is no evidence favoring multi-domain CCT over single domain. It's true that the authors, appropriately, do not argue that there is a difference – that multi-domain CCT is more efficacious – but without further comment, a less than careful reader might come away with that impression.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

1. The meta-analysis needs to be registered on Prospero to ensure that the protocol & results are available to other researchers for replication and updating. It is the only thing missing from the Prisma checklist.

Thank you for the suggestion to register the meta-analysis on Prospero (https://www.crd.york.ac.uk/prospero/). We agree that the protocol should be available to other researchers, however unfortunately at the stage of registration of our protocol, data extraction was complete and the study was therefore ineligible to be registered. As stated on the Prospero website:

“Reviews that have completed data extraction are not eligible for inclusion in PROSPERO. The aim of the register is to capture information at the design stage.”

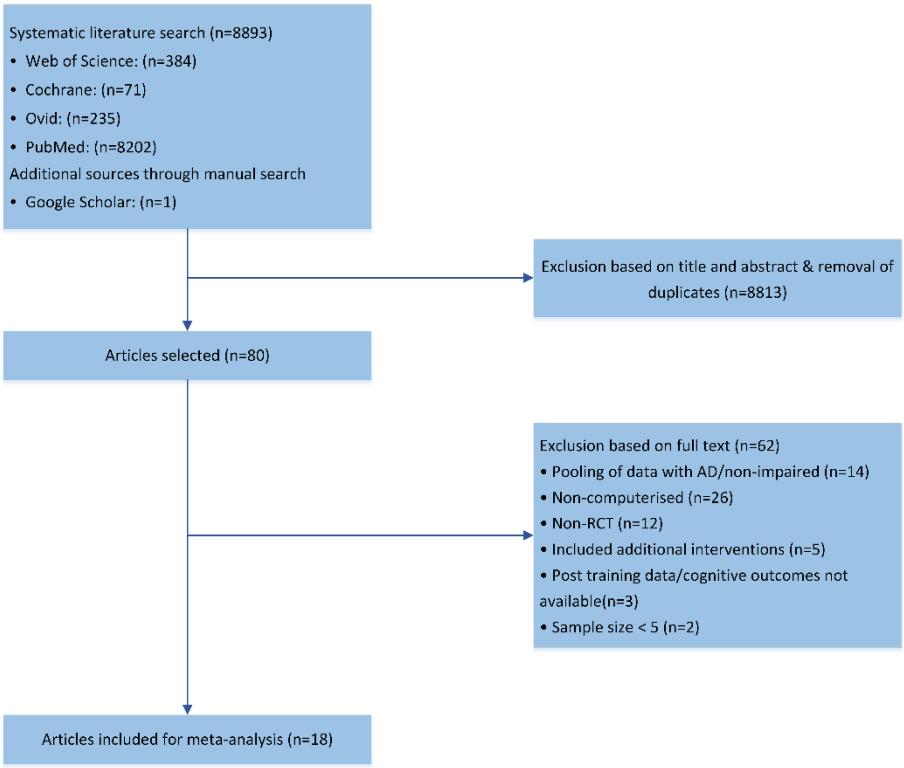
We have added the lack of registration as a study limitation in lines 361-365.

Lines 361-365. Another limitation of the present meta-analysis is the lack of registration on Prospero. The registration could ensure that the protocol & results are available to other researchers for replication and updating. However unfortunately at the stage of registration of our protocol data extraction was complete and the study was therefore ineligible to be registered on Prospero.

2. The authors should perform their search on PubMed to ensure that no studies were missed as there are different subject matter articles on PubMed compared to Web of Science.

Thank you for this suggestion to perform the search on PubMed. We have added the search results of PubMed until Jan 2018, consistent with when other databases were searched. This is detailed in lines 148-149 and Figure 1.

Lines 148-149. A literature search was completed during January 2018 of four online literature databases and trial registers: PubMed, Embase, Web of Science and Cochrane library.



3. Table 1 is excellent but too comprehensive to be included in the main body of the paper - perhaps shorten the summary for each study and move the majority of the table to an online supplement.

Thank you for your kind comments about the contents in table 1. We have shortened it and moved the majority of the table to the supplementary table 3.

Table 1. Characteristics of studies using computerised cognitive training in persons with MCI

Author and Year	CCT Group N, age, education	Control Group N, age, education	CCT type	Total hours
Barban et al 2016	N = 46, Age = 74.4 (5.7), Edu = 9 (4.3)	N = 60, Age = 72.9 (6.0), Edu = 11 (4.7)	Multi domain	24
Ciarmiello et al 2015	N = 15, Age = 71.2 (7.7), Edu = 9.3 (3.0)	N = 15, Age = 72.0 (7.1), Edu = 7.8 (2.6)	Multi domain	24
Djabelkhjr et al 2017	N = 10, Age = 75.2 (6.4), Edu = 60.0% of college level	N = 10, Age = 78.2 (7.0), Edu = 44.4% of college level	Multi domain	18
Fiatarone et al 2014	N = 24, Age = >55, Edu = n/s	N = 27, Age = >55, Edu = n/s	Multi domain	80
Finn & McDonald 2011	N = 8, Age = 69.0 (7.7), Edu = 13.3 (2.2)	N = 8, Age = 76.4 (6.5), Edu = 12.0 (2.8)	Multi domain	25
Finn & McDonald 2015	N = 12, Age = 72.8 (5.7), Edu = 13.8 (3.0)	N = 12, Age = 75.1 (7.5), Edu = 13.7 (2.8)	Memory	n/s
Gagnon & Belleville 2012	N = 12, Age = 67.0 (7.8), Edu = 15.0 (4.6)	N = 12, Age = 68.4 (6.0), Edu = 13.1 (5.7)	Attentional control	6
Gooding et al 2016 study 1	N = 31, Age = 75.6 (8.8), Edu = 15.1 (2.6)	N = 10, Age = 75.6 (8.8), Edu = 15.1 (2.6)	Multi domain	30
Gooding et al 2016 study 2	N = 23, Age = 75.6 (8.8), Edu = 15.1 (2.6)	N = 10, Age = 75.6 (8.8), Edu = 15.1 (2.6)	Multi domain	30
Hagovska et al 2016	N = 40, Age = 68.0 (4.4), Edu = 75% of secondary education	N = 40, Age = 65.9 (6.2),Edu = 70% of secondary education	Multi domain	10

Author and Year	CCT Group N, age, education	Control Group N, age, education	CCT type	Total hours
Han et al 2017	N = 23, Age = 73.7 (4.8), Edu = 13.5 (3.2)	N = 20, Age = 74.5 (6.4), Edu = 12.7 (3.7)	Memory	4
Herrera et al 2012	N = 11, Age = 75.1 (2.0), Edu = 46% of secondary school or more	N = 11, Age = 78.2 (1.4), Edu = 63% of secondary school or more	Multi domain	24
Hughes et al 2014	N = 10, Age = 78.5 (7.1), Edu = 13.8 (2.4)	N = 10, Age = 76.2 (4.3), Edu = 13.1 (1.9)	Multi domain	36
Hyer et al 2016	N = 34, Age = 75.1 (7.4), Edu = 70% secondary	N = 34, Age = 75.2 (7.8), Edu = 66% secondary	Working memory	16.7
Lin et al 2016	N = 10, Age = 72.9 (8.2), Edu = 90.0% of college level	N = 11, Age = 73.1 (9.6), Edu = 54.5% of college level	Processing speed	24
Rosen et al 2011	N = 6, Age = 70.7 (10.6), Edu = 16.7 (0.8)	N = 6, Age = 78.0 (7.9), Edu = 18.3 (1.5)	Processing speed	36
Rozzini et al 2007	N = 15, Age = 63-78,Edu = n/s	N = 22, Age = 63-78, Edu = n/s	Multi domain	60
Savulich et al 2017	N = 21, Age = 75.2 (7.4), Edu = 15.9 (1.3) (Age left school)	N = 21, Age = 76.9 (8.3) Edu = 16.0 (2.1) (Age left school)	Memory	8

Notes: MMSE: Mini Mental State Examination, n/s: not stated

Reviewer 2:

The paper is a systematic review and meta-analysis of studies assessing the efficacy of computerized cognitive training for late middle aged and elderly people with mild cognitive impairment. The paper is well-written and carefully conducted, and their methods and analysis are effectively described with good and appropriate detail.

We thank the reviewer for their kind comments.

However, given the recent review by Hill, Mowszowski, Naismith et al. (Computerized cognitive training in older adults with mild cognitive impairments or Dementia: a systematic review and meta-analysis. American Journal of Psychiatry, 2017; 174:329-340), a question that needs to be more fully addressed is what is added contribution this paper provides. The authors cite the earlier review and other related ones, but they need to show what their report adds to the question of CCT efficacy beyond what has already been published.

We have made the following additions to emphasise the contribution this study makes to the existing literature in Lines 124-135 and Lines 317-327.

Lines 124-135. Systematic reviews and meta-analysis of cognitive interventions in MCI have reported mixed results, and when exploring the effect of cognitive training in MCI have largely not distinguished between studies evaluating computerised and non-computerised training. This makes it difficult to draw conclusions specifically on the efficacy of CCT in MCI. For example, a systematic review by Ge et al summarised the findings of CCT studies among people with MCI, however no meta-analyses were performed and the review included non-randomized controlled studies, studies that combined CCT with other interventions, and studies not using Petersen's core MCI diagnosis criteria making it challenging to draw rigorous conclusions. A previous meta-analysis by Hill et al specifically explored the effectiveness of CCT in MCI on cognition and behavioural outcomes, however, the field is progressing rapidly, as highlighted by Ge et al's observation that 42% of the studies in their review were published between 2016 and 2017, and further relevant studies have been published subsequently. Another more recently published meta-analysis by Gates et al only included studies where the intervention period lasted for more than 12 weeks and excluded a significant number of studies with shorter training duration. Thus, it is necessary to conduct an updated meta-analysis to include more recent articles and all intervention durations.

Lines 317-327. The present meta-analyses updated the literature search and added eight new studies compared with the previous study conducted by Hill et al. The present findings are largely in keeping with the results of Hill et al that demonstrated positive effect sizes for global cognition (g=0.38, 95% CI=[0.14–0.62]), memory (g=0.42, 95% CI =[0.21, 0.63]), working memory (g=0.74 , 95% CI =[0.32, 1.15]) and executive function (g=0.20, 95% CI=[-0.05, 0.44]). However, our results are in contrast with the results reported by Gates et al which found that there were no clear effects of CCT on cognition for people with MCI. Methodological reasons for this inconsistency may be that Gates et al only included studies with a minimum intervention period of 12 weeks and included a broader range of participants at risk of cognitive decline. As a result, much fewer studies (eight) met their eligibility criteria, of which two studies did not require a strict MCI diagnosis [47 48] and one used self or informant–reported cognitive complaints[56].

Other very minor concerns are itemized below.

1. In the Abstract and the Introduction, the authors write that their objective is to “determine the efficacy” (line 37) and that it “investigates the efficacy” (line 133) of CCT for MCI, but as this is not a direct study of CCT, but a review of other studies, this language might be revised.

Thank you for this suggestion. We have changed the “determine/investigates the efficacy” to “determine/evaluate the effect” (Title, Lines 24, 142).

2. Line 124-5 states “Critical analysis of research using for MCI should reveal insight into any effective components of CCT….” This was somewhat unclear. Perhaps what is meant is, “Ideally, critical analysis…would reveal insight into specific components…”?

Thank you for your helpful suggestion about the unclear statement. We have changed it to “Ideally, critical analysis of research using CCT for MCI would reveal insight into which specific components of CCT are necessary for it to be effective” (Lines 117-120).

3. Line 245 discusses the effect sizes found in studies of global cognition and each of the cognitive domains using active control groups versus those with non-active controls. While the stats and significance of the CCT versus control effect sizes are provided, it would also be useful to test the differences between the effect sizes. For example, beginning on line 254, it is noted that for studies of memory using active controls, the “effect size was statistically significant and larger than that of trials with passive control groups” (the latter was not significant). Might these different effect sizes be tested for statistical significance?

Thank you for your suggestion about statistically comparing effect sizes. We have further performed a meta-regression, as detailed in Lines 193-195, 254-257, 263-266, 284-287, 296-299, and 303-304.

Line 193-195. We also performed subgroup analysis and meta-regression using the "metafor" program in R (https://www.R-project.org/), for example we compared the effectiveness of single and multi-domain training.

Lines 254-257. The effect size across active-controlled trials (n=7, g=0.23, 95% CI [-0.05, 0.51], I^2=27%) was smaller than that of trials with non-active control groups (n=4, g=0.31, 95% CI [-0.06, 0.68], I^2=0%) (see supplementary figure 3-4.), but was not statistically significantly different (z = -0.11, p = 0.91).

Lines 263-266. The effect size across active-controlled trials (n=8, g=0.36, 95% CI [0.11, 0.61], I^2=52%) was larger than that of trials with passive control groups (n=5, g=0.20, 95% CI [-0.14, 0.54], I^2=43%) (see supplementary figure 6-7.), but was not statistically significantly different (z = -0.32, p = 0.75). However, there was moderate heterogeneity across studies in both analyses.

Lines 284-287. The effect size across active-controlled trials (n=7, g=0.13, 95% CI [-0.08, 0.35], I^2=20%) was smaller than for the non-active control groups (n=4, g=0.32, 95% CI [-0.23, 0.87],

I^2=74%) (see supplementary figure 8-9.), but was not statistically significantly different (z = 0.95, p = 0.35).

Lines 296-299. Our subgroup analyses and meta-regression suggested that there is no difference between multi-domain CCT and single-domain CCT (z = 0.09, p = 0.93), although the former had a significant effect (g = 0.30, 95% CI (0.08, 0.53)) while the latter was non-significant (g = 0.31, 95% CI (-0.19, 0.81)) (see supplementary figure 10-11).

Lines 303-304. We did not perform a meta-regression for training dose because fewer than ten studies were included.

4. In another subgroup analysis – multi-domain versus single domain CCT – the g statistic was significant for multi-domain CCT, but not for single domain CCT. However, the g statistic – was nominally larger for the single domain CCT (.31) than the multi-domain CCT (.30). This suggests that the difference in significance is probably a function of the reduced power in the single domain studies. A test comparing these two g-statistics would surely be non-significant and would show that, at least based on the available studies to date, there is no evidence favoring multi-domain CCT over single domain. It is true that the authors, appropriately, do not argue that there is a difference – that multi-domain CCT is more efficacious – but without further comment, a less than the careful reader might come away with that impression.

Thank you for your comments about the subgroup analysis and we agree that it is important that this is not misleading. We have changed the description in Lines 296-299:

Lines 296-299: Our subgroup analyses and meta-regression suggested that there is no difference between multi-domain CCT and single-domain CCT (z = 0.09, p = 0.93), although the former had a significant effect (g = 0.30, 95% CI (0.08, 0.53)) while the latter was non-significant (g = 0.31, 95% CI (-0.19, 0.81)) (see supplementary figure 10-11).

VERSION 2 – REVIEW

REVIEWER	Isabel Elaine Allen University of California San Francisco USA
REVIEW RETURNED	04-Apr-2019

GENERAL COMMENTS	Well done, all reviewer comments answered.

REVIEWER	Jeremy M. Silverman Icahn School of Medicine at Mount Sinai, USA James J. Peters Veterans Affairs Medical Center, USA
REVIEW RETURNED	15-Apr-2019

GENERAL COMMENTS	The authors have been highly responsive to the prior reviews,. The paper is both a careful review of CCT studies for MCI and a useful

	discussion of many methodological issues associated with clinical trials in the area. They have convincingly shown that their paper does indeed make a valuable contribution to the existing literature in this area.