Review Article
Systematic Review and Meta-Analysis of Clinically Relevant Executive Functions Tests Performance after COVID-19

Boris B. Velichkovsky,1,2 Anna Y. Razvaliaeva,3 Alena A. Khlebnikova,2 Piruza A. Manukyan,1,2 Vladimir N. Kasatkin,1 and Artem V. Barmin4

1Research Institute for Brain Development and Peak Performance, Peoples’ Friendship University of Russia, Moscow 117198, Russia
2Lomonosov Moscow State University, Moscow 125009, Russia
3Institute of Psychology, Russian Academy of Sciences, Moscow 129366, Russia
4Cognitive Foundations of Communication Laboratory, Moscow State Linguistic University, Moscow 119034, Russia

Correspondence should be addressed to Anna Y. Razvaliaeva; anna.razvaliaeva@gmail.com

Received 8 November 2022; Revised 12 January 2023; Accepted 27 January 2023; Published 9 February 2023

Academic Editor: Luigi Trojano

Copyright © 2023 Boris B. Velichkovsky et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is widely known that COVID-19 has a number of prolonged effects on general health, wellbeing, and cognitive functioning. However, studies using differentiated performance measures of cognitive functions are still not widely spread making it hard to assess the exact functions that get impaired. Taking into account the similarities between post-COVID ‘brain fog’ and chemofog, we hypothesized that executive functions (EF) would be impaired. Literature search yielded six studies with 14 effect sizes of interest; pooled effect size was small to medium ($d = -0.35$). Combined with a narrative synthesis of six studies without a comparison group, these results show that EF get impaired after COVID-19; although, in most cases the impairment is transient and does not seem to be severe. These results specify the picture of ‘brain fog’ and may help to discover its mechanisms and ways of helping people with long COVID.

1. Introduction

Since it emerged in December 2019, COVID-19 has caused a global healthcare crisis and posed a number of problems. Concerns about its prolonged impact on survivors have arisen, corroborated by their frequent reports of fatigue, cognitive impairments (also dubbed ‘brain fog’), shortness of breath, sleep disturbances, depression, anxiety and post-traumatic stress disorder, impaired or distorted sense of smell and taste, muscle pain, and headaches [1, 2]. These symptoms have a negative impact on the quality of life, mental health, and employment, can last for months after recovery and emerge even in non-hospitalized or asymptomatic patients [1, 3]. Prolonged effect of COVID-19 could also be observed in escalation in severe mental illnesses, such as psychosis [4].

It is hypothesized that cognitive impairment after COVID-19 is linked to the impact of the disease on neurological, immune, and vascular systems [5, 6]. Patients in more severe condition (e.g., requiring ventilation) may develop multiple diffuse lesions in the white matter [7]. Global cognitive impairment after COVID-19 has also been linked to volume changes in the white matter, especially in the frontal lobes [8]. These and other neurological complications may be contributed to the effects of neuroinflammation, neurotoxicity, or hypoxia [5]. It is worth noting that hypoxemia can be considered as the cause of psychiatric long-term consequences of COVID-19, such as an escalation in psychosis [9]. According to the recent study [10] among the possible causes of the so-called ‘Cognitive COVID’ (acute changes during COVID-19 and after it) are systemic inflammation, multisystem inflammatory syndrome and ventilation, psychological strain caused by pandemic and lockdowns, direct neurotropism, and subsequent brain damage. However, the exact mechanisms behind the cognitive sequelae have not been discovered yet, and one of the
possible reasons for that is a lack of studies focusing on specific cognitive functions. Executive functions (EF) could be a good focus point because their impairments are associated with conditions closely resembling post-COVID ‘brain fog’.

EF are broadly defined as a prerequisite for performing complex tasks where automatic behaviors are maladaptive, for example, when rules change, new distractors emerge or risks require assessment. Key EF include inhibition, cognitive flexibility, selective attention, and working memory; they are closely related to fluid intelligence, problem solving and decision-making skills, and quality of life [11]. A series of studies by Miyake and Friedman illustrates that EF contain common components, but shifting and updating have specificity that cannot be explained by the common factor [12]. The latent common factor uniting EF can explain the methodological problem with validity and ‘task purity’, when measures designed to assess a single EF can be related to other complex executive and non-executive mental processes [13]. Specificity, on the other hand, can explain the use of different strategies in solving common problems [12, 14].

Evidence points to disruptions in EF due to somatic illnesses, especially in chemobrain, a condition prevalent in cancer patients after chemotherapy and characterized by cognitive decline and difficulties in everyday tasks [15]. Cancer survivors show significantly lower scores on inhibition, switching, and planning tasks, compared with normative scores [16]. These changes are related to decreases in white matter volume and functional changes [17]. Links between EF (visuospatial learning, lexical access, and set shifting), cytokines, and helper and cytotoxic T-cell count have also been discovered in chronic fatigue syndrome research [18]. Inhibition, working memory, and set shifting are also impaired in adolescents with type 1 diabetes [19]. Given that neuroinflammation plays an important role both in chemobrain and brain fog [20], we could expect similar deficits in EF in people with post-COVID syndrome.

Mental distress also has a negative impact on EF. Acute stress leads to worsening accuracy in task management, planning, and coding measures [21], as well as working memory and task flexibility impairments [22]. The effect of acute stress on inhibition is not straightforward: while it impairs cognitive inhibition, response inhibition slightly increases, which signifies greater ability to control one’s behavior [22]. Inhibitory control, planning, flexibility, decision-making, and sustained attention scores are lower in students with depression and anxiety symptoms [23]. Given the high prevalence of stress-related mental disorders in COVID-19 survivors and their correlation with cognitive deficits [24], this is another factor that could contribute to deficits in EF.

The current study was carried out in order to determine to what extent EF get impaired after recovery from COVID-19. The research question was formulated in accordance with population–exposure–comparator–outcome (PECO) formula as follows: How do EF change in people after COVID-19 in comparison with healthy controls?

2. Materials and Methods

2.1. Eligibility Criteria. Studies were included if they: (a) described people with confirmed COVID-19, (b) and compared them with healthy controls on (c) objective cognitive measures of EF. Studies were excluded if they: (a) used global measures of cognitive functioning or self-report measures, did not assess particular EF and (b) focused on patients with other chronic illnesses or cognitive dysfunction preceding COVID-19. Studies without a healthy control group were excluded from meta-analysis, but if they used objective tests of EF, they were included in the narrative review.

2.2. Information Sources. We conducted searches in Scopus, Web of Science, and PubMed citation databases. Medical and psychological preprint databases (PsyArXiv, medRxiv, bioRxiv, and Research Square) were included in accordance with the guidelines on decreasing publication bias. The latter also helped to gain more papers given the novelty of the field and the stringency of the eligibility criteria.

2.3. Search Strategy and Selection Process. Initial searches were conducted in November–December 2021. Keywords related to SARS-CoV-2, cognitive deficit, and EF were used. In July–August 2022 citation bases were searched again in an attempt to discover more papers that fit eligibility criteria.

Search results were uploaded to the CADIMA online tool for systematic reviews [25]. After duplicate removal, titles and abstracts were screened independently by three reviewers (AKh, PM, and AR).

2.4. Data Collection. The following data items were collected from the studies:

(i) Study information: authors and year of publication.
(ii) Demographic information: number of participants, sex, age, country of residence, and education.
(iii) COVID-specific information: time since illness and illness severity (number of hospitalized patients in COVID + group and number of those who required intensive care and/or ventilation).
(iv) Results of measuring EF: means and standard deviations, names of tests, and/or scales.

Working memory and attention measures were not included in the current meta-analysis as they were analyzed and described in detail elsewhere [26].

2.5. Risk of Bias Assessment. To assess study quality we used the Johanna Briggs Institute Critical Appraisal Checklist for analytical cross sectional studies [27]. Studies were analyzed independently by three reviewers; consensus on any points of disagreement was reached during a group discussion.

2.6. Pooling Effect Sizes Scales. To measure effect sizes, standardized mean differences (Cohen’s d) were calculated on the basis of group means, standard deviations, and sample sizes reported in the original studies. If the study reported medians and ranges [28], they were converted to means and standard deviations according to formulas provided by Hozo et al. [29].
We used random-effect models to pool effect sizes (as opposed to fixed-effect models that operate on the assumption of a single true effect size). This approach is effective when the original studies employ different samples (e.g., from different countries), data gathering techniques and measures; that is, when the assumption of a homogeneous population cannot be upheld [30].

Given that a lot of studies use more than one test to measure EF and report numerous effect sizes, we decided to use a multilevel approach when pooling them. It can reduce researcher bias stemming from picking out a single effect from each study.

Heterogeneity was expected to be high based on variability in measures, study designs, and sample characteristics. To measure it $I^2$ statistic was used, and the value of 75% or more was considered indicative of high heterogeneity.

Sensitivity analysis was performed using leave-one-out approach, when the model was retested after removing a single observed outcome at a time. This approach can show which study contributed to the pooled effect size the most. Publication bias was tested by Egger’s regression test. In case it could not be calculated directly we modelled it by including standard errors as a moderator in the model.

Data were analyzed in the R version 4.1.2 [31]. The following packages were used:

(i) esc ver. 0.5.1 [32]—to compute effect sizes (Cohen’s $d$);

(ii) metaphor ver. 3.0.2 [33]—to pool effect sizes, perform sensitivity analysis, assess publication bias, and draw plots.

3. Results

3.1. Study Selection. The results of the selection process are shown in Figure 1. Twelve relevant papers were identified as a result of the search. Six studies were only included in the narrative synthesis because they reported EF scores only in post-COVID group. Thus, six papers were included in the meta-analysis. Participants, measures, and overall bias scores for the studies are shown in Table 1.

In case of Wild et al. [34] patients with acute COVID were included in the sample, which made it necessary to recalculate overall effect sizes excluding this group of participants.

All studies were of acceptable quality. There were several concerns raised in quality assessment. Online-based study design applied by Wild et al. [34] and Zhao et al. [35] could lead to self-selection of participants who did not have laboratory-confirmed COVID-19, but self-diagnosed. These studies also did not report the country of residence of the sample making it harder to determine possible cultural input into the results. Guo et al. [36] included participants without laboratory-confirmed COVID-19 and did not report illness severity (% of patients who had required hospitalization). Mattioli et al. included participants as early as 12 days after recovery [28].

3.2. Measures Used in the Studies. The following measures of EF were used in the studies: Trail Making Task-B, Stroop Color-Word Interference test, Tower of London task, and Wisconsin Card Sorting Task. The tasks are described in Table 2. Although all these tasks involve multiple EF, Stroop Color-Word Interference test is considered a measure of inhibition and Tower of London measures planning and reasoning. Wisconsin Card Sorting Task and Trail Making Task-B are used in brain injury testing (the former in particular is considered a frontal lobe task) and measure set shifting, task switching, and updating.

3.3. Results of Individual Studies. A total of 14 effect sizes were collected from the studies (Table 3). Effect sizes (Cohen’s $d$) ranged from −0.80 to 0.10; a half of them were statistically significant.

Studies not included in meta-analysis also reported a decrease in EF in post-COVID patients. Hampshire et al. [37] discovered larger decreases in planning (measured by the Tower of London task) for people who had been hospitalized and put on ventilation compared with those who did not exhibit or had only mild respiratory symptoms. Effect size for this subgroup of patients was small to medium and ranged from −0.36 for all cases (including those who suspected COVID) to −0.42 for those with a positive COVID-19 test [37].

Silva et al. [38] reported an abnormally low performance on the Trail Making Task-B for 40% of the sample of 87 non-hospitalized individuals. They also found a correlation ($r = 0.30$) between performance on this task and white matter impairments in inferior longitudinal fasciculus (an association tract that connects occipital and temporal lobes).

Miskowiak et al. [39] compared patients’ ($n = 29$) performance on Trail Making Task-B to sex-, age-, and education-adjusted norms revealing a large significant decrease in cognitive flexibility ($d = −0.81$). In another study by the same group of authors hospitalized patients ($n = 29$) performed worse on the same task than non-hospitalized ones ($n = 19$); $d = −0.78$ [40].

Larsson et al. [41] found no differences between ICU-admitted patients and those who had less severe symptoms on Trail Making Task-B ($d = −0.06$); however, older participants (65+ years old) performed this task significantly slower than younger ones ($d = −0.54$).

Surprisingly, comparing scores of patients after COVID-19 ($n = 31$, approximately 6 months after positive test) with population norms on Color-Word Interference task and Trail Making Task-B, Dressing et al. [42] did not find any participants with scores below mean value.

3.4. Effect Sizes Synthesis. The pooled standardized mean difference between post-COVID-19 participants and healthy controls was −0.35; 95% CI [−0.59, −0.11], $p < 0.05$. Figure 2 presents the forest plot for the model.

Variance testing on different levels showed that Level 3, or between-study heterogeneity ($I^2 = 58.79$) was larger than Level 2, or within-study heterogeneity ($I^2 = 27.97$). Sampling error (Level 1) variance was 13.24. Overall heterogeneity was large ($I^2 = 86.76$), possibly due to different measures used in the studies, and different study designs (online and offline).
Sensitivity analysis was carried out in two ways. First, effect sizes were removed one by one. Pooled effect size varied from $-0.40$ to $-0.32$; the largest value was reached when the results of Tower of London test from Wild et al. [34] were removed. The removal of the same test from Poletti et al. [43] on the contrary impacted the value in the opposite direction. A more traditional leave-one-out analysis was also performed when studies were removed on the whole. In this case pooled effect size varied from $-0.43$ (after six effect sizes from Guo et al. [36] were removed) to $-0.26$ (after three effect sizes from Crivelli et al. [44] were removed).

Publication bias could not be assessed directly (due to the use of multilevel modeling). The model was rerun, and standard error was included as a moderator. According to the results publication bias can be considered insignificant ($b = -1.84$, 95% CI $[-5.28, 1.59]$, SE $= 1.58$, $t = -1.17$, $p = .27$). However, they should be interpreted with caution due to the interdependencies between effect sizes coming from the same study.

Figure 1: PRISMA flowchart for study identification and selection.
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Location</th>
<th>nF</th>
<th>Age (years)</th>
<th>Edu</th>
<th>nF</th>
<th>Age (years)</th>
<th>Edu</th>
<th>Time elapsed</th>
<th>Severity</th>
<th>EF measures</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crivelli et al., 2022</td>
<td>Cross-sectional</td>
<td>Argentina</td>
<td>20</td>
<td>57 [46–64]</td>
<td>17</td>
<td>50 [43–63]</td>
<td>17</td>
<td>[15–18]</td>
<td>142 ± 75.9 days</td>
<td>14 (31%) were hospitalized</td>
<td>Trail Making Task-B and Wisconsin Card Sorting Test</td>
<td>100</td>
</tr>
<tr>
<td>Guo et al., 2021a</td>
<td>Cross-sectional</td>
<td>UK 137 + 130; North America 24 + 33</td>
<td>118</td>
<td>[18 ± 61]</td>
<td>181</td>
<td>130</td>
<td>[18 ± 61]</td>
<td>63% PSE</td>
<td>3 ± 31 weeks</td>
<td>NA</td>
<td>Wisconsin Card Sorting Test</td>
<td>93</td>
</tr>
<tr>
<td>Mattioli et al., 2021</td>
<td>Cross-sectional</td>
<td>Italy</td>
<td>72</td>
<td>40.57 ± 11.79</td>
<td>312</td>
<td>52.63 ± 8.81</td>
<td>12.94 ± 3.76</td>
<td>118 mild-moderate; 2 in intensive care unit (ICU)</td>
<td>BACS Tower of London</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poletti et al., 2021</td>
<td>Cross-sectional</td>
<td>Italy</td>
<td>72</td>
<td>40.57 ± 11.79</td>
<td>312</td>
<td>52.63 ± 8.81</td>
<td>12.94 ± 3.76</td>
<td>118 mild-moderate; 2 in intensive care unit (ICU)</td>
<td>BACS Tower of London</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild et al., 2021a</td>
<td>Prospective cohort</td>
<td>NA</td>
<td>7832</td>
<td>42.76 ± 14.44</td>
<td>478</td>
<td>43.41 ± 13.17</td>
<td>82.47% post-sec</td>
<td>15 asymptomatic; 403 mild COVID; 67 hospitalized, 17 in ICU</td>
<td>Tower of London; modified Stroop task ('double trouble')</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhao et al., 2021a</td>
<td>Cross-sectional</td>
<td>NA</td>
<td>25</td>
<td>25.03 ± 7.33</td>
<td>36</td>
<td>25</td>
<td>27.71 ± 8.51</td>
<td>167.3 ± 127.9 days</td>
<td>100% mild cases, no hospitalization</td>
<td>Tower of London (version of the original task)</td>
<td>93</td>
<td></td>
</tr>
</tbody>
</table>

Note: a—non-peer-reviewed preprint; nF—number of women; Edu—education in years or % of sample with a degree; PSE—post-secondary education; NA—information not available; EF—executive functions; BACS—brief assessment of cognition in schizophrenia. Data format—mean ± SD [range].
4. Discussion

People who recover from COVID-19 often report cognitive dysfunction as one of the long-lasting symptoms. Although memory difficulties and poor attention and concentration are more often cited in reference to these symptoms, objective measures show that EF (such as inhibition, updating, and set shifting) also get impaired. Effects range from small to large, and generally decreases in performance are related to advanced age and illness severity. These results corroborate...
earlier systematic analyses [45, 46] and further validate their points by pooling effect sizes from studies with healthy control groups.

Current findings can be translated into clinical practice through diagnostics of cognitive dysfunctions and creating cognitive rehabilitation programs involving the cognitropic pharmacology administration, cognitive training, and the use of transcranial direct current stimulation (tDCS). Cognitropic pharmacology administration aims to enhance cognitive functions contributing to one’s longevity [47]. Cognitive training aims to maintain or improve cognitive functions through practicing cognitive tasks [48]. Cognitive training as part of COVID-19 rehabilitation may consist of practicing working memory tasks, taking into account the possible improvement of other functions, such as response inhibition [49]. Unfortunately, the effectiveness of cognitive training is still limited [50]. The use of tDCS in patients with Alzheimer’s disease and Parkinson’s disease represents an effective cognitive rehabilitation strategy [51].

A possible direction of cognitive rehabilitation is the restoration of unspecific cognitive resources after COVID-19. According to Kahneman [52], higher-order cognitive tasks require the use of unspecific, shared, and exhaustible cognitive ‘resources’. Specific effects on the brain, such as brain fog can be the cause of cognitive resources depletion. Thus, the study of the distribution of cognitive resources in patients with brain fog can play a significant role in the creation of cognitive rehabilitation programs.

The pooled effect size \( (d = -0.35) \) is close to the result yielded by a meta-analysis of chemobrain in cancer survivors \( (d = -0.26) \) [16], thus providing another link between these conditions apart from immune and neurological changes [20]. Given that both brain fog and chemobrain are related to a noticeable loss in the quality of life, rehabilitation programs for long COVID could benefit from including EF training already used for cancer survivors.

Effect sizes of a similar magnitude were also yielded for short-term memory in people with long COVID [26]. It brings forth the question on the structure and relationships between different impaired cognitive functions. EF could be at the root of short-term memory deficits, or both could be equally impaired due to cognitive exhaustion caused by fatigue and inflammatory processes.

### 5. Limitations

First, not enough studies fit the criteria. Most studies found by the initial search either utilized a cohort design (no control group) and/or used self-report measures and global test of cognitive functioning (no assessment of EF). Second, there was an inconsistency in the results of the studies yielded by the search: some showed large and significant effect sizes, whereas others found no relationships between the same variables. This could be attributed to the large variation between sample characteristics and measures of EF. Third, it is important to remember that about two thirds of the data were obtained from studies conducted more than 5 years ago, which may limit the applicability of the findings to the current population.
of COVID-19 survivors do not report cognitive problems and studies used in the current review selected participants with cognitive complaints and other post-COVID symptoms (e.g., fatigue).

6. Conclusions

In this meta-analysis, we have provided evidence of the negative effect of prolonged COVID-19 consequences for executive cognitive functions. Self-reported post-COVID cognitive impairments are usually associated with memory and attention. However, deficits in inhibition, switching, and planning can also be observed in COVID-19 survivors with long-lasting symptoms. This effect was observed in studies comparing post-COVID participants (without previous cognitive dysfunction) and healthy controls on objective measurements of specific EF. The results of the study can specify the picture of brain fog and show its similarity to chemobrain in terms of EF deficiency. The results can contribute to the debates on the matter of post-COVID cognitive impairments for discovering ways of helping people with prolonged consequences of COVID-19.

Data Availability

The data supporting this systematic review and meta-analysis are from previously reported studies, which have been cited. The processed data (effect sizes) are included within the article (Table 3).

Conflicts of Interest

The author(s) declare(s) that they have no conflicts of interest.

Acknowledgments

This publication has been supported by the RUDN University Scientific Projects Grant System [project number 212200-2-000].

References


