Research Report

A two-week running intervention reduces symptoms related to depression and increases hippocampal volume in young adults

Andreas Fink*, Karl Koschutnig, Thomas Zussner, Corinna M. Perchtold-Stefan, Christian Rominger, Mathias Benedek and Ilona Papousek

Institute of Psychology, University of Graz, Austria

Abstract

This study examined the effects of a two-week running intervention on depressive symptoms and structural changes of different subfields of the hippocampus in young adults from the general population. The intervention was realized in small groups of participants in a mostly forested area and was organized into seven units of about 60 min each. The study design included two intervention groups which were tested at three time points and which received the intervention time-delayed: The first group between the first and the second time point, and the second group between the second and the third time point (waiting control group). At each test session, magnetic resonance imaging (MRI) was performed and symptoms related to depression were measured by means of the Center for Epidemiological Studies Depression (CES-D) Scale. Results revealed a significant reduction of CES-D scores after the running intervention. The intervention also resulted in significant increases in the volume of the hippocampus, and reductions of CES-D scores right after the intervention were associated with increases in hippocampal volume. These findings add important new evidence on the beneficial role of aerobic exercise on depressive symptoms and related structural alterations of the hippocampus.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Physical activity is associated with beneficial effects on various cognitive functions including attention, inhibition, cognitive flexibility, or memory-related demands (Masley et al., 2009; Smith et al., 2013). Likewise, there is a wealth of studies demonstrating positive effects of physical activity on academic performance in school-aged children (e.g., de Greeff et al., 2018; Donnelly et al., 2016; Sardinha et al., 2014). Being
physically active has been also found to reduce symptoms of schizophrenia and depressive symptoms in people suffering from mental illness (e.g., Rosenbaum et al., 2014). A similar picture of findings has been revealed in non-clinical populations (Rebar et al., 2015). In examining dose and domain of physical activity to reduce symptoms of depression in adults, Teychenne et al. (2008) concluded that even relatively low physical activity to reduce symptoms of depression in adults, populations (Rebar et al., 2015). In examining dose and domain of physical activity which are necessary for unfolding beneficial effects on depression are not clarified yet (Teychenne et al., 2008). Nevertheless, evidence seems to be strong enough to support the promising role of physical activity in improving mental health (Dinas et al., 2011; Mammen & Faulkner, 2013).

In a similar vein, neuroscience studies revealed positive effects of physical activity on both functional and structural characteristics of the brain, especially in samples of school-aged children (Donnelly et al., 2016) and in older adults (Erickson et al., 2011). For example, Erickson et al. (2014) reviewed studies investigating the relationship between cardiorespiratory fitness, physical activity, exercise interventions and gray matter volume in late adulthood. On the basis of about 20 studies since 2003, they concluded that higher cardiorespiratory fitness levels were associated with larger gray matter volume in regions closely linked with cognitive functions such as the prefrontal cortex or the hippocampus (see also Ratey & Loehr, 2011). Especially the hippocampus has been identified as being particularly sensitive to physical activity (Erickson et al., 2011; Firth et al., 2018; Gujral et al., 2017; Killgore et al., 2013; Rendeiro & Rhodes, 2018). For example, a meta-analysis of Firth et al. (2018) included 14 studies involving different aerobic exercise interventions (e.g., stationary cycling interventions, walking-based intervention, etc.) mainly in clinical samples and in samples of older adults. They found that aerobic exercise was associated with significant positive effects on left hippocampal volume in comparison to control conditions. The study of Thomas et al. (2016) was the only study of this meta-analysis which focused on young to middle aged adults. The authors tested a 6-week stationary cycling intervention and found an increase in the volume of the anterior hippocampus after the intervention, which returned to baseline six weeks afterwards during which time the participants were no longer engaged in regular aerobic exercise. Similar evidence for early to middle-aged adults has been revealed by Nauer et al. (2020) who found increases in the volume of the hippocampal head after a 12-week moderate-intensity exercise intervention designed to increase cardiorespiratory fitness.

Structural alterations of the hippocampus are also reported in people suffering from depression (Campbell & MacQueen, 2004; Chen et al., 2020; Espinoza Oyarce et al., 2020; Nogovitsyn et al., 2020). For example, a comprehensive meta-analytic study by Gujral et al. (2017) reported that the volume of the hippocampus was about 5% smaller in depression. In a recent study, Nogovitsyn et al. (2020) performed magnetic resonance imaging (MRI) scans in people suffering from major depressive disorder and in healthy controls. This study revealed that the baseline volume of the hippocampal tail was significantly smaller in people suffering from major depression compared to healthy controls. Interestingly, larger baseline volumes of the hippocampal tail were positively associated with the remission status after treatment, indicating that the volume of this brain structure could constitute a prognostic biomarker for the treatment outcome in people suffering from major depressive disorder. Postmortem studies even indicated that the hippocampal volume and the number of neurons and glial cells is reduced in people with a history of depression and schizophrenia relative to controls by an amount of 20%–35% (Chen et al., 2020). Hence, given that depression is linked with lower hippocampal volumes, and given that exercise interventions have been found to modulate structural characteristics of the hippocampus (see Gujral et al., 2017, who specifically examined the overlap of structural brain alterations related to depression and exercise related changes on brain structure), it seems reasonable to assume that exercising may ameliorate depressive symptoms, possibly via changes in functional and structural characteristics of the hippocampus. Neuronal growth factors such as the brain-derived neurotrophic factor (BDNF) are often discussed as underlying biochemical mechanisms implicated in neural changes of the hippocampus (von Bohlen & Halbach, 2018), especially also in the context of physical exercise (Erickson et al., 2012; Liu & Nusslock, 2018; Smith et al., 2013).

While controlled randomized trials examining the effects of physical activity interventions on depression and structural brain changes are generally rather rare, most of the available studies focused on older adults (Erickson et al., 2011; Hillman et al., 2008). Erickson et al. (2019), for instance, recently identified noticeable gaps of research in some populations, especially in adolescents, and young/middle-aged adults. For this reason, the present study examined the effects of a two-week running intervention on depressive symptoms and structural changes of the hippocampus in a sample of young adults from the general population (mostly university students). Along with the fact that evidence concerning structural brain changes and depressive symptoms in young adults is limited, focusing on young adulthood avoids potential confounding effects of medical conditions on brain structure in later adulthood. In addition, it has been found that individuals who regularly exercise at earlier stages of their lifetime may be better protected against cognitive decline in later adulthood (Ratey & Loehr, 2011).

The running intervention of this study was organized into seven units of about 50–60 min each, conducted over a time period of two weeks. The running sessions were realized in small groups of participants, supervised by at least one experimenter. The standardized running route lead through a mostly forested area at a local recreation area and was about five kilometers long. The design of this study included two groups of participants who were tested at three time points of assessment. At each test session, MRI scans were performed and depressive symptoms were measured by means of a self-assessment scale. The participants were randomly assigned to a first running group receiving the running intervention between the first (t1) and the second time point of assessment (t2), or to a second running group which received the running intervention between t2 and the third time point.
of assessment (t3), thereby acting as waiting control group. The period between t1 and t2 allows to assess whether the first running group shows the expected intervention-related decreases in depressive symptoms and increases in hippocampal volume. Critically, the third time point of assessment allows to examine whether the second running group likewise shows the expected intervention-related decreases in depressive symptoms and increases in hippocampal volume right after the intervention (i.e., from t2 to t3). In addition, it also facilitates the assessment of whether any intervention-related effects remain stable over time (relevant for the first intervention group).

2. Method

This study examined the effects of a running intervention on symptoms related to depression and structural changes of different subfields of the hippocampus in young adults from the general population. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. No part of the study procedures and study analyses was pre-registered. The measures (symptoms related to depression and hippocampal volume) are described in this section, along with the intervention design. The data that support the findings of this study are available under https://openneuro.org/datasets/ds003799/versions/1.0.0. MRI Quality Control reports of each image are included in the derivatives/mriqc folder on openneuro. Sample size was determined on the basis of similar neuroscientific training studies in this field (e.g., Fink et al., 2015). In the Flow diagram in Fig. 1, we itemize drop outs of participants over the study period. During the recruitment (especially prior to MRI assessment) participants were carefully screened for potential exclusion criteria (e.g., metallic implants, cardiac pacemaker, piercings, tattoos, etc.); they signed a questionnaire to guarantee that they fulfil all required safety requirements.

2.1. Participants

A sample of 68 participants (mostly university students) was recruited for this study. As itemized in Fig. 1, 11 participants did not participate in the MRI assessment sessions, resulting in a sample of 57 participants who consented to participate in the running intervention and in the MRI assessment sessions. Out of this sample, 48 participants (27 women) completed all three MRI scans and psychometric assessments and participated in the running intervention (9 missed one or two scanning session/s, see Fig. 1). Their age ranged between 19 and 33 years (M = 23.00, SD = 2.82); all participants were right-handed (except for one who indicated ambidexterity), with no history of neurological or psychiatric disorders. We primarily recruited rather unathletic people showing no or only low regular engagement in sports activities. This was motivated by the presumption that in people who do not exercise regularly, potential effects of a brief running intervention on structural brain changes and corresponding changes of symptoms related to depression would be more pronounced (as compared to people who do a lot of sports). Participants indicated to exercise about half an hour per week (M = .53; SD = 1.2). They received course credit and were paid for their participation in the MRI assessment sessions. Written informed consent was obtained. The study was approved by the authorized ethics committee.

2.2. Study design and procedure

Participants were randomly assigned to two intervention groups, which received the running intervention time-delayed (see Fig. 1). The first group included 21 participants (11 women) and performed the running intervention between the first (t1) and the second test session (t2), while the second...
group (wait list control group, n = 27, 16 women) received the training between t2 and the third test session (t3). The distribution of sex was not significantly different across both groups (Chi² (1) = .23, P = .63). Also, there were no significant group differences in age ([t(46) = .72, P = .48] and regarding the amount of exercising per week ([t(46) = .57, P = .57]).

In using this design (three test sessions and two training groups), we expect the first intervention group to show significant intervention-related changes in brain structure and related affective functions from the first to the second test session, while the second group should show no substantial changes during this time. The latter group is expected to exhibit intervention-related changes from t2 to t3, i.e., right after the running intervention. The data of t3 of the first intervention group also allow examining whether any training effects remain stable over time, while the data from the second group at t2 provide a test of potential changes that are unrelated to training (e.g., re-test effects). Another advantage of this design could be seen in the important ethical issue that all participants could be recruited for this study in the very same way (in promising the participation in a supervised running intervention) and the participants received the same intervention (which is often not the case in conventional control group designs).

2.3. The intervention

The standardized running route lead through a mostly forested area at a local recreation area next to the university. The route was about five kilometers long, with a height difference of about 100 m. Given the fact that the participants of this study were beginners with no or little experience in running, the running intervention could be considered as a strong increase in physical activity in these participants. Accordingly, the team of researchers took great care to maintain or increase study commitment and motivation to the best possible extent. During recruitment, it was clearly communicated that is important for the success of the study that no sessions (running units and MRI appointments) are missed. The research team was in regular contact with the participants, and they continuously received reminders. During the run itself, participants were running in a moderately intense, self-paced tempo. The running units also included periods of warming up and cooling down. Due to individual differences in running speed, larger groups of participants were often split up into smaller subgroups. In such cases, one experimenter was running with the faster participants, while another experimenter formed the taillight. Together with periods of warming up and cooling down, the total duration of a single running intervention was about 60 min.

There were several appointments for the running intervention. Participants were free to choose among these possible running appointments. Every participant was required to participate seven times in this running intervention over a time period of two weeks.

2.4. Psychometric tests

At each time point of assessment (t1, t2, and t3), psychometric tests for the assessment of different affective (Positive and Negative Affect Schedule, PANAS; Krohne et al., 1996; State-Trait Anxiety Inventory, STAI; Laux et al., 1981) and cognitive functions (Test for Creative Thinking Drawing Production; Urban & Jellen, 1995) were administered. Legal copyright restrictions prevent public archiving of these psychometric tests, which can be obtained from the copyright holders in the cited references.

The focus of this study was to test intervention-related changes of symptoms related to depression, which were assessed by means of the German version of the Center for Epidemiological Studies Depression Scale (CES-D; Hautzinger et al., 2012). The CES-D is a brief self-report scale designed to assess symptoms related to depression in the general population (Radloff, 1977), and might thus be especially suited in this study with young adults from the general population. Importantly, the CES-D scale does not provide any diagnosis of clinical manifestations of depressive symptomatology, rather it is considered as “a first-stage screener to target respondents with depressive symptoms for more in-depth clinical assessment” (Vilagut et al., 2016, p. 13).

The CES-D comprises 20 items covering depressive symptoms in the emotional, motivational, cognitive, somatic, and motoric domain. Participants had to indicate the extent to which the given conditions apply to them in the last week on a four-point scale ranging from “infrequent” (0) to “the most time” (3). In the test manual, a cut-off score of 22 is suggested for further diagnostic clarification (Hautzinger et al., 2012). At the first time point of assessment, the participants showed a mean CES-D score of 12.85 (SD = 7.57; Skewness = 1.01, SE of skewness = .34), very similar to the data of representative samples from the general population (Hautzinger et al., 2012).

2.5. MRI assessment and analysis

2.5.1. MRI data acquisition

MRI data was acquired on a 3.0 T Tim-Dot Magnetom Skyra system with syngo MR E11 (Siemens Medical Systems, Erlangen, Germany) using a 32-channel head coil. T1 measurements were obtained using a 3D mprage-sequence derived from the commercial VE11C package with the following parameters: 224 sagittal slices [.88 mm slice thickness], FOV read 224 mm, TR 2200 msec, TE 2.18 msec, base resolution 256 [acquisition matrix 256x256, .88 mm iso-voxel, no partial fourier], flip angle 8 deg, phase encoding direction AP, phase oversampling 15%, PAT 2 [GRAPPA, accel. factor PE 2], single shot, interleaved, bandwidth 210 Hz/Px. The measurement time for the T1-mprage sequence was 5 min 51 sec.

2.5.2. MRI data processing

After removal of facial features (https://github.com/poldracklab/pydeface/tree/master/pydeface), all T1-weighted images were reviewed with the MRI Quality Control tool (MRIQC; Esteban et al., 2017). MRIQC reports of each image are included in the derivatives/mriqc folder on openneuro (https://openneuro.org/datasets/ds003799/versions/1.0.0). MRIQC reports of each image are included in the derivatives/mriqc folder on openneuro (https://openneuro.org/datasets/ds003799/versions/1.0.0). FreeSurfer is a freely available software for the processing of neuroimaging data, including automatized, standardized, and
well-proven algorithms and routines for image processing for both cross-sectional and longitudinal research designs (a full documentation of this software can be found at https://surfer.nmr.mgh.harvard.edu). Freesurfer also includes processing tools for the automated segmentation of subfields of the hippocampus, which is of central interest in this study.

Longitudinal image data were processed cross-sectionally for all the time points following the default Freesurfer workflow including multiple time points per subject (‘recon-all’). In a first step, all sets of two time points (e.g., participant-1, t1 - t3) were calculated independently with the longitudinal stream in FreeSurfer (‘recon-all -long’) followed by the computation of individual intra subject templates (Reuter et al., 2012). Afterwards, the results of longitudinal calculation were compared in native space.

In a second step we performed the longitudinal hippocampal subfield segmentation (“segment”) based on the work of Iglesias et al. (2015) and Iglesias et al. (2016). The Freesurfer tool is based on an atlas of statistical probability, generated with ultra-high resolution ex vivo MRI data, for automated hippocampal subfield allocation on subject level. The method also segments the nuclei of the amygdala, further increasing segmentation accuracy, since overlapping between structures is minimized (Saygin, Kliemann, et al., 2017). For further post-processing, the hippocampal subfields were aggregated into the Head, Body and Tail (“HBT”) of the hippocampus.

2.6 Statistical analyses

To investigate the effects of the running intervention on symptoms related to depression (CES-D scores), a General Linear Model (GLM) for repeated measures with the TIME point of assessment as within-subjects factor (t1, t2, t3) and the experimental GROUP as between-subjects factor (first vs second intervention group) was conducted. Separate repeated measures GLMs for the volumes of the head, the body, and the tail of the hippocampus were performed, including the factors TIME (t1, t2, t3), HEMISPHERE (left vs right) and experimental GROUP (group 1 and 2). Significant intervention effects are reflected in significant interactions involving the factors TIME and GROUP, i.e., both intervention groups are expected to show changes in CES-D scores and hippocampal volume right after the running intervention. In case of significant interaction effects, follow-up paired t-tests were computed for both intervention groups separately.

To further test the effects of the running intervention, we investigated intervention-related changes of hippocampal volume as a function of changes of symptoms related to depression. For this reason, we computed the difference in CES-D scores between the pre-test (immediately before the running intervention) and the post-test (right after the intervention) according to the formula: CES-D post-test – CES-D pre-test. Thus, negative values indicate a decrease in CES-D scores, while positive values are indicative of increases in symptomatology. In the majority of cases we found decreases in CES-D scores right after the running intervention (CES-D scores from t2 – t1 in group 1: M (SD) = –4.71 (6.24), Range: –20 to 6; and t3 – t2 in group 2: M (SD) = –2.11 (5.37), Range: –14 to 8). These pre-test versus post-test changes in CES-D scores were then used as continuous predictor variables in statistical analyses. Specifically, for each experimental group and for each hippocampal subfield (head, body, tail, and whole hippocampus), separate GLMs for repeated measures involving the within-subjects factors HEMISPHERE (left vs right) and TIME (t1, t2, t3), and the continuous predictor “CES-D change” were computed. These analyses should reveal information whether or to which extent intervention-related changes in CES-D scores were associated with corresponding changes in hippocampal volume: Regarding group 1, we expect that intervention-related decreases in the CES-D score (at t2 as compared to t1) are linked with corresponding increases in hippocampal volume during that time. In group 2, decreases in CES-D scores after the intervention (at t3 as compared to t2) should be associated with increases in hippocampal volume at t3 versus t2.

To illustrate significant effects involving the continuous between-subjects factor (i.e., pre-test vs post-test changes in CES-D scores), predicted hippocampal volumes were calculated for CES-D changes one standard deviation below and one standard deviation above the sample mean using standard regression analysis.

In case of violations of sphericity assumptions, degrees of freedom were Greenhouse-Geisser corrected. The significance level was $P < .05$ in all statistical analyses. For the GLM repeated measures analyses, estimates of effect sizes are given in terms of partial eta-squared measures ($\eta^2_p$). For the t-Tests, Cohen’s d values are reported.

3. Results

3.1 Intervention effects on symptoms related to depression (CES-D scores)

The 2 (GROUP) x 3 (TIME) repeated measurements GLM for the CES-D revealed a significant interaction between TIME and GROUP ($F(2, 92) = 5.02, P < .01, \eta^2_p = .10$), along with a significant TIME effect ($F(2, 92) = 3.57, P = .03, \eta^2_p = .07$). As shown in Fig. 2, both groups, especially the first running group showed reductions in CES-D scores right after the running intervention. Interestingly, participants of this group showed re-increases in CES-D scores from t2 to t3, during which time they were no longer engaged in the running intervention. Subsequent paired t-tests computed separately for both running groups indicated substantial reductions of CES-D scores right after the intervention in both groups (group 1 from t1 to t2: $t(20) = 3.46, P < .01, d = .76$; group 2 from t2 to t3: $t(26) = 2.04, P = .051, d = .39$). The re-increase in group 1 was not significant (from t2 to t3: $t(20) = –1.88, P = .075, d = -.41$).

3.2 Intervention effects on the hippocampal volume (head, body, tail)

The repeated measures GLM revealed a significant interaction between TIME and GROUP in the hippocampal tail ($F(2, 92) = 4.90, P < .01, \eta^2_p = .10$). As shown in Fig. 3, the first running group showed volume increases right after the intervention, and a weak decline from t2 to t3. Follow-up paired t-tests revealed a nearly significant increase in the first running group from t1 to t2 ($t(20) = –2.07, P = .051, d = –...
and a significant decline in the second group from t1 to t3 ($t_{(26)} = 2.89, P < .01, d = .56$).

No significant intervention-related effects regarding the volumes of the head and the body of the hippocampus were found (Head: $F(2, 92) = 1.67, P = .19, \eta^2 P = .03$; TIME x HEMISPHERE x GROUP: $F(2, 92) = .50, P = .61, \eta^2 P = .01$; Body: TIME x GROUP: $F(1.75, 80.44) = 1.83, P = .17, \eta^2 P = .04$; TIME x HEMISPHERE x GROUP: $F(2, 92) = 1.14, P = .32, \eta^2 P = .02$).

The relevant interaction between TIME x GROUP also reached significance for the total hippocampal volume ($F(2, 92) = 3.17, P = .047, \eta^2 P = .06$, along with a significant HEMISPHERE effect ($F(1,46) = 6.94, P = .01, \eta^2 P = .13$; more volume in the right hippocampus). The pattern of this interaction was similar to that observed for the hippocampal tail: The first training group showed a continuous increase in the total hippocampal volume over the time points of assessment, especially between t1 and t2 (significant was the increase from t1 to t3, $t_{(20)} = -2.11, P = .048, d = -0.46$). The second training group showed no significant differences between the three time points of assessments.

---

**Fig. 2** – CES-D scores at t1, t2, and t3 separately for both running groups. Group 1 received the running intervention between t1 and t2, group 2 between t2 and t3. Both groups showed reductions of CES-D scores right after the running intervention. Error bars indicate ± 1 SE.

---

**Fig. 3** – Volume of the hippocampal tail at t1, t2, and t3 for both running groups. Group 1 showed increases in the volume of the hippocampal tail right after the intervention. Error bars indicate ± 1 SE.
3.3. Intervention-related changes of hippocampal volume as a function of changes of CES-D scores

The repeated measures GLM with TIME (t1, t2, t3) and HEMISPHERE (left, right) as within-subjects factors and the CES-D change right after the intervention as continuous between subjects factor in the first running group (CES-D at t2 – CES-D at t1) revealed a significant interaction between HEMISPHERE, TIME, and CES-D change for the volume of the hippocampal tail, F(1.45, 27.63) = 4.63, P = .03, \( \eta^2_p = .20 \). The pattern of this interaction indicated that participants showing stronger reductions of CES-D scores right after the running intervention (at t2 versus t1) exhibited stronger increases of the volume of the right hippocampal tail during that time (see Fig. 4). However, the interpretation of this interaction is somewhat complicated by the fact that this effect seems to be driven primarily by two participants who showed comparatively strong reductions in CES-D scores and strong increases in hippocampal volume (see Fig. 4). Regarding the head and the body of the hippocampus, no significant effects involving TIME and CES-D change were found.

For the second running group, the repeated measures GLM with TIME (t1, t2, t3) and HEMISPHERE (left, right) as within-subjects factors and the CES-D change right after the intervention as continuous between subjects factor (CES-D at t3 – CES-D at t2) revealed a significant interaction between TIME and CES-D change for the volume of the hippocampal body (F(2, 50) = 5.50, P < .01, \( \eta^2_p = .18 \)). Fig. 5 reveals that this interaction effect was driven by the changes from t2 to t3, i.e., immediately before and right after the running intervention: The stronger the decreases in CES-D scores right after the training, the stronger were the increases in the volume of the hippocampal body. The interaction between TIME and CES-D change was likewise significant for the hippocampal head in this group (F(1.54, 38.46) = 3.73, P = .044, \( \eta^2_p = .13 \). Again, there was an inverse association between changes in CES-D scores and changes in hippocampal volume (see Fig. 5). In the hippocampal tail, no significant effects involving TIME and CES-D change were found. As illustrated in Fig. 5, for the total hippocampal volume this interaction again reached significance (F(2, 50) = 5.97, P < .01, \( \eta^2_p = .19 \)).

4. Discussion

A moderate running intervention realized in seven 60 min units over a time period of two weeks resulted in a significant reduction of symptoms related to depression in a sample of young adults from the general population. Strikingly, the running intervention also modulated the structure of the hippocampus. To the very best of our knowledge, this is the first study demonstrating exercise-related changes of brain structure and related affective functions after such a short intervention. A recent systematic review of eight meta-
analyses comprising a total of 134 individual studies revealed a moderate effect of exercise interventions on reducing depressive symptoms in the general population (Hu et al., 2020). The duration of the included interventions ranged from four weeks to about one year, that is at least twice as long as the intervention in this study. Of even greater importance in the present study is the fact that already two weeks of running modulated the structure of the hippocampus, which is known as critical component of the neural network implicated in depression (e.g., Gujral et al., 2017; Kandola et al., 2019; Nogovitsyn et al., 2020; Roddy et al., 2019).

While the first running group generally showed increases in the volume of the hippocampal tail, in the second running group no general changes of the hippocampal volume after the intervention were evident (post hoc t-tests for the hippocampal tail even revealed a continuous decline from t1 to t3 in

![Graphs showing changes in hippocampal volume and CES-D scores](image-url)
this group). This was paralleled by the behavioral finding that the reduction of CES-D scores in the second running group was somewhat smaller than in the first intervention group (reduction of $M = -2.1$ vs $M = -4.7$ points in the CES-D for the second vs first running group, respectively). A closer look on the training conditions of the second group may reveal some possible explanations why this group was somewhat less successful. 1) A violent storm during the training period of this group reduced the number of training units from 7 to 6 in some participants of this group; 2) as a consequence of the storm, parts of the original running route were closed by official order and the route had to be slightly changed into a more urban environment in these participants; and 3) the training period of this group was closer to the examination week at the end of the university semester, possibly acting as additional stressor counteracting intervention effects. Though inadvertently, these somewhat more unfavorable training conditions in the second running group seem to hint at issues related to dose and response. It seems that the “optimal” dose of intervention was slightly too low in this group to modulate the structure of the hippocampus.

Critically and importantly, however, there was an association between volumetric increases of the hippocampus and reductions of CES-D scores in both groups. As depicted in Figs. 4 and 5, participants showing stronger reductions in CES-D scores right after the running intervention exhibited stronger increases in the volume of the hippocampal tail (group 1), and in the head, body, and whole hippocampus (group 2). This significant association between reductions of symptoms related to depression and volumetric increases of the hippocampus further corroborates the crucial role of this brain structure in depression. However, it is not clear from this study which subfields of the hippocampus are most sensitive to the running intervention. The first group showed general intervention-related increases in the hippocampal tail, which was further modulated by the amount of CES-D change in this group (see Fig. 4). While this finding is well in line with very recent research reports showing that hippocampal tail volume predicts depression status and remission (Maller et al., 2018; Nogovitsyn et al., 2020), other studies showed volume increases in anterior portions of the hippocampus after exercising (e.g., Erickson et al., 2011; Thomas et al., 2016). These findings seem to suggest that, despite some possible overlap of the underlying neuronal networks, exercise and depression-related effects on the structure of the hippocampus may be also driven by different biological mechanisms. Irrespective of the location of the effects, these studies clearly suggest that it is important to study subfields of the hippocampus in relation to exercise and depression, rather than the structure of the hippocampus as a whole.

A strength of this study design is that it also allows to assess whether intervention effects remain stable over time. As the data of the first running group suggested, there was a trend towards a re-increase of depressive symptoms in the absence of the running intervention. The volume of the hippocampus remained stable from t2 to t3, though at a purely descriptive level, a weak decrease of volume was observed. One of the rare studies in this field, which employed additional follow-up assessments (Thomas et al., 2016) showed that the intervention-related increase in the volume of the hippocampus is temporary and returns to baseline after additional six weeks without aerobic exercise. Together with the findings of this study it hence seems that a continuous and regular engagement in physical activity is necessary to maintain brain structure and related affective functions.

We hope that the findings of this study stimulate new studies employing exercise interventions of different mode and length, to address one of the most challenging question in this context—the question of dose–response. The application of state-of-the-art objective measures of fitness, psychometric measures for the assessment of different cognitive/affective functions, and modern brain imaging methods in well-controlled longitudinal intervention studies may provide excellent prerequisites to achieve a more fine-grained understanding of how exercise may support brain and affective functions and mental health. Of course, the specific psychological and biological mechanisms of how physical exercise unfolds its antidepressant effects needs to be clarified in future research. Kandola et al. (2019) recently summarized several important psychological and biological factors in this context. These include changes in neuroplasticity, inflammatory processes, the endocrine system, and oxidative stress, along with psychosocial factors such as self-esteem, social support and self-efficacy. This nicely exemplifies that this field is highly interdisciplinary, requiring a fine-grained interplay of different scientific disciplines.

Among the most important limitations of this study is the small sample size. This study required beginners in running to participate in a two-weeks running intervention and to participate in three MRI scans at three different time points. In addition, available scan time was limited. Together with missing MRI appointments (see Fig. 1), this resulted in a comparatively low number of participants, especially in group 1, which complicates the scope of the findings (cf. Fig. 4). Therefore, it would be exciting to see whether the findings of this study can be replicated in larger and more powerful samples of participants.

This study also challenges the crucial question of dose and response. The participants of this study indicated that they were previously exercising only about half an hour per week, on average. In the intervention of this study they were required to participate in seven running units (~5 km, ~100 m difference in altitude) over a time period of two weeks. This could be considered as a considerable increase in physical activity, and it may well be the case that the running intervention constituted a high intensity exercise for the study participants. A comprehensive literature review of Cooper et al. (2016) showed that high-intensity intermittent exercise is associated with highly beneficial effects on indicators of cardiometabolic health and cognitive functions/academic achievements in young people. However, the specific kind and dose of exercise, which is most likely to unfold brain changes and corresponding changes in cognitive and affective functions and indicators of mental health is among the most challenging research questions in this field.

As outlined in the method section, processing of the structural T1-weighted datasets was performed by using the longitudinal Freesurfer pipeline (http://surfer.nmr.mgh.harvard.edu/, version 6.0.0). Since the available scan time was limited, we decided to include only one structural
sequence (T1-weighted MPRAGE) in our sequence protocol. Therefore, hippocampal subfield-segmentation was calculated using only T1-weighted scans. The pipeline allows for the use of an additional, structural scan for hippocampal subfield-segmentation. Therefore, in addition to the T1-weighting, either a T2-weighting with increased resolution, or ideally susceptibility weighted images (SWI) could further increase segmentation reliability.

Author contributions

Andreas Fink: Conceptualization; Formal analysis; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Writing - original draft. Karl Koschutnig: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Visualization; Writing - review & editing. Thomas Zussner: Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Visualization; Writing - review & editing. Corinna M. Perchtl-Stefan: Conceptualization; Methodology; Writing - review & editing. Christian Rominger: Conceptualization; Methodology; Writing - review & editing. Mathias Benedek: Conceptualization; Methodology; Writing - review & editing. Ilona Papousek: Conceptualization; Methodology; Writing - review & editing.

Open practices

The study in this article earned an Open Data badge for transparent practices. Data for this study can be found at https://openneuro.org/datasets/ds003799/versions/1.0.0.

Funding information

Expense allowance for participating in the MRI sessions was provided by the Faculty of Natural Sciences of the University of Graz.

Data availability statement

The data that support the findings of this study are available under https://openneuro.org/datasets/ds003799/versions/1.0.0.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors wish to thank Katharina Lenz, Michaela Lingenhel, Anna Mayr, Theresa Rumpf, Melanie Saceric and Laura Sammer for their great engagement in realizing the running interventions and in collecting the data.

REFERENCES

Campbell, S., & MacQueen, G. (2004). The role of the hippocampus in the pathophysiology of major depression. Journal of Psychiatry & Neuroscience, 29(6), 417–426. PMID: 15644983; PMCID: PMC524959.

Chen, F., Bertelsen, A. B., Holm, I. E., Nyengaard, J. R., Rosenberg, R., & Dorph-Petersen, K. A. (2020). Hippocampal volume and cell number in depression, schizophrenia, and suicide subjects. Brain Research, 1727, 146546. https://doi.org/10.1016/j.brainres.2019.146546

Cooper, S. B., Dring, K. J., & Nevill, M. E. (2016). High-intensity intermittent exercise: Effect on young people’s cardiometabolic health and cognition. Current Sports Medicine Reports, 15(4), 245–251. https://doi.org/10.1249/JSR.0000000000000273

De Greeff, J. W., Bosker, R. J., Oosterlaan, J., Visscher, C., & Hartmann, E. (2018). Effects of physical activity on executive functions, attention and academic performance in preadolescent children: A meta-analysis. Journal of Science and Medicine in Sport, 21, 501–507. https://doi.org/10.1016/j.jsams.2017.09.595

Dinas, P. C., Koutedakis, Y., & Flouris, A. D. (2011). Effects of exercise and physical activity on depression. Irish Journal of Medical Science, 180(2), 319–325. https://doi.org/10.1007/s11845-010-0633-9

Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., Lambourne, K., & Szabo-Reed, A. N. (2016). Physical activity, fitness, cognitive function, and academic achievement in children: A systematic review. Medicine and Science in Sports and Exercise, 48, 1197–1222. https://doi.org/10.1249/MSS.0000000000000901

Erickson, K. I., Hillman, C., Stillman, C. M., Ballard, R. M., Bloodgood, B., Conroy, D. E., Macko, R., Marquez, D. X., Petruzzello, S. J., Powell, K. E., & For 2018 Physical Activity Guidelines Advisory Committee. (2019). Physical activity, cognition, and brain outcomes: A review of the 2018 physical activity guidelines. Medicine and Science in Sports and Exercise, 51(6), 1242–1251. https://doi.org/10.1249/MSS.0000000000001936

Erickson, K. I., Leckie, R. L., & Weinstein, A. M. (2014). Physical activity, fitness, and gray matter volume. Neurobiology of Aging, 35(Suppl 2), S20–S28. https://doi.org/10.1016/j.neurobiaging.2014.03.034

Erickson, K. I., Miller, D. L., & Roecklein, K. A. (2012). The aging hippocampus: Interactions between exercise, depression, and BDNF. The Neuroscientist: a Review Journal Bringing Neurobiology, Neurology and Psychiatry, 18(1), 82–97. https://doi.org/10.1177/1073858410397054

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., Kim, J. S., Heo, S., Alves, H., White, S. M., Wojcicki, T. R., Mailey, E., Vieira, V. J., Martin, S. A., Pence, B. D., Woods, J. A., McAuley, E., & Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. Proceedings of the National Academy of Sciences of the United States of America, 108(7), 3017–3022. https://doi.org/10.1073/pnas.1015950108

Espinoza Oyarce, D. A., Shaw, M. E., Alateeq, K., & Cherbuin, N. (2020). Volumetric brain differences in clinical depression in association with anxiety: A systematic review with meta-analysis. Journal of Psychiatry & Neuroscience, 45(6), 406–429. https://doi.org/10.1503/jpn.190156

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). Mriqc: Advancing the automatic prediction of image quality in MRI from unseen sites. Plos One, 12(9), Article e0184661. https://doi.org/10.1371/journal.pone.0184661
Smith, P. J., Potter, G. G., McLaren, M. E., & Blumenthal, J. A. (2013). Impact of aerobic exercise on neurobehavioral outcomes. Mental Health and Physical Activity, 6(3), 139–153. https://doi.org/10.1016/j.mhpa.2013.06.008

Teychenne, M., Ball, K., & Salmon, J. (2008). Physical activity and likelihood of depression in adults: A review. Preventive medicine, 46(5), 397–411. https://doi.org/10.1016/j.ypmed.2008.01.009

Thomas, A. G., Dennis, A., Rawlings, N. B., Stagg, C. J., Matthews, L., Morris, M., Kolind, S. H., Foxley, S., Jenkinson, M., Nichols, T. E., Dawes, H., Bandettini, P. A., & Johansen-Berg, H. (2016). Multi-modal characterization of rapid anterior hippocampal volume increase associated with aerobic exercise. Neuroimage, 131, 162–170. https://doi.org/10.1016/j.neuroimage.2015.10.090

Urban, K. K., & Jellen, H. G. (1995). Test zum schöpferischen denken-zeichnerisch. Frankfurt: Swets.

Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): A systematic review with meta-analysis. Plos One, 11(5), Article e0155431. https://doi.org/10.1371/journal.pone.0155431

von Bohlen und Halbach, O., & von Bohlen und Halbach, V. (2018). BDNF effects on dendritic spine morphology and hippocampal function. Cell and Tissue Research, 373, 729–741. https://doi.org/10.1007/s00441-017-2782-x