

Testing 'mommy brain': no evidence of altered memory-related brain function across the transition to motherhood

Sophie van 't Hof

s.r.vanthof@amsterdamumc.nl

Amsterdam UMC <https://orcid.org/0000-0003-1758-8966>

Willeke Menks

Donders Institute for Brain, Cognition, and Behaviour, Radboud University and Radboud University Medical Center

Marlieke van Kesteren

Vrije Universiteit

Eveline A. Crone

Erasmus University Rotterdam

Elseline Hoekzema

Amsterdam University Medical Center <https://orcid.org/0000-0001-5814-8647>

Article

Keywords:

Posted Date: October 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10699834/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Testing 'mommy brain': no evidence of altered memory-related brain function across the transition to motherhood

Authors

van 't Hof, S.R.^{1*} s.r.vanthof@amsterdamumc.nl
https://orcid.org/0000-0003-1758-8966
Menks, W.M.² willeke.menks@donders.ru.nl
van Kesteren, M.T.R.³ marliekeevk@gmail.com
Crone, E.A.^{4,5} crone@essb.eur.nl
https://orcid.org/0000-0002-7508-6078
Hoekzema, E.^{1,4*} e.a.hoekzema@amsterdamumc.nl
https://orcid.org/0000-0001-5814-8647

1. Pregnancy Brain Lab, Amsterdam University Medical Center (AUMC), location University of Amsterdam, Department of Psychiatry, Amsterdam Neuroscience, Amsterdam Reproduction and Development, Amsterdam, the Netherlands

2. Donders Institute for Brain, Cognition, and Behaviour, Radboud University and Radboud University Medical Center, Nijmegen, the Netherlands

3. Institute of Brain and Behavior Amsterdam, Vrije Universiteit, 1081 BT, Amsterdam, The Netherlands

4. Brain and Development Research Center, Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands

5. Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, the Netherlands

* corresponding author

Abstract

Pregnancy induces structural brain changes, but effects on task-based brain function remain unexamined. Here we investigated verbal memory-related brain function across this transition using a pre-conception longitudinal fMRI design in 110 women (first-time and second-time mothers, and nulliparous controls), and assessed memory and subjective complaints in an independent replication cohort (n=144). No significant group differences in encoding-related brain activity or objective memory performance were observed. Both maternal groups reported significant, persistent increases in subjective memory complaints, absent in controls and linked to change in depressive symptoms and perceived stress rather than pregnancy itself. These findings show no evidence of decline in memory-related brain function or performance, and suggest that psychological wellbeing, not pregnancy or parity, shapes both subjective complaints challenging the neurobiological basis of a 'mommy brain'.

1 Introduction

2 Pregnancy is a significant transformative life phase, marked by profound changes in
3 hormonal and physiological processes. Research by our and other groups has shown that
4 pregnancy renders substantial changes in brain structure¹⁻⁶, also in women becoming mothers
5 for the second-time³, involving strong volumetric reductions within frontoparietal networks
6 that support higher-order cognitive functions. Building on these findings, we recently showed
7 that becoming a mother is also associated with changes in human brain function, specifically
8 an increase in temporal coherence in the Default Mode Network².

9 However, despite these advances in understanding pregnancy-related structural and resting-
10 state brain changes, no study has examined task-evoked neural activity. Such data are
11 essential to understand whether the transition to motherhood alters specific brain functions
12 during active cognitive processing. Still, in the absence of direct functional evidence,
13 observed reductions in grey matter volume have often been interpreted as evidence of
14 cognitive decline, commonly referred to as 'mommy brain' or 'baby brain'⁷⁻⁹ in popular
15 discourse. Yet, self-reported as well as objective behavioral evidence for memory changes
16 remain mixed^{8,10-13}.

17 One of the most-studied types of memory in this context is verbal memory. Behavioral
18 findings vary, with some studies reporting deficits in verbal paired associates learning and
19 free recall during late pregnancy^{14,15}, whereas recognition memory appears consistently
20 preserved across studies^{9,16,17}. Longitudinal studies with pre-conception baselines have
21 yielded inconsistent results: while one study reported selective deficits in verbal recall
22 alongside hormonal changes across pregnancy¹⁴, the largest prospective cohort study found
23 no objective decline in verbal memory across pregnancy and motherhood¹⁸. Together, these
24 findings suggest subtle strategic retrieval differences rather than a fundamental impairment in
25 memory encoding. Furthermore, where deficits are observed, studies rarely account for
26 concomitant changes in for example sleep, stress, and depression, factors known to
27 independently impair cognitive performance during the peripartum period^{8,19,20}, making it
28 difficult to attribute observed effects to pregnancy or early postpartum itself.

29
30 Although behavioral results are inconsistent, 'mommy brain' remains a cultural stereotype
31 widely discussed in popular culture^{21,22}. Critically, interpreting these structural changes as
32 evidence of cognitive decline may be misguided. Existing research points to a framework of

1 adaptive neural reorganization, in which pregnancy-related brain changes are concentrated in
2 regions supporting social cognition, threat detection, and motivational processing, functions
3 directly relevant to caregiving^{1,23-25}. In addition, structural and resting-state changes across
4 pregnancy have been linked to prepartum and postpartum mother-infant bonding behavior^{1,2},
5 resting-state connectivity in postpartum mothers has been associated with their maternal
6 behavior^{26,27}, and pregnant women show enhanced encoding and retention of parenting-
7 relevant stimuli compared to neutral stimuli¹⁰. This evidence thus lends support to
8 evolutionary perspective suggesting that maternal brain plasticity serves functions relevant to
9 infant care rather than representing a loss of capacity²⁸. This adaptive framing does not
10 preclude the possibility that prioritizing neural resources toward caregiving-relevant
11 functions carries trade-offs for other, less directly prioritized domains; rather, it predicts that
12 any such changes should be functionally selective rather than reflecting broad cognitive
13 decline. Under this framework, memory-related brain activity during encoding might be
14 expected to remain largely stable across the transition to motherhood, with any observed
15 changes more likely reflecting contextual factors often accompanying this transition such as
16 sleep disruption, distress, and depression than pregnancy-specific neural decline^{8,19,20}.

17
18 We aimed to investigate the neural basis of ‘mommy brain’ and examined whether becoming
19 a mother is associated with changes in verbal memory-related brain functioning by means of
20 a pre-conception cohort study involving first-time mothers, second-time mothers, and
21 nulliparous control women and four time points: before conception, during late pregnancy, at
22 the early and the late postpartum periods. Changes in verbal memory-related brain activity
23 were measured with a well-established memory encoding paradigm^{29,30}, with successful
24 memory encoding quantified via a composite memory score used as a parametric modulator
25 at the first level.

26
27 First, we examined whether the expected pattern of encoding-related brain activity was
28 replicated in our sample. Then, to examine whether the transition to first- or second-time
29 motherhood is associated with different changes in memory-related brain functioning, we
30 compared the changes in brain function from pre-conception until the early postpartum period
31 in first-time mothers (n = 40), second-time mothers (n = 30), and control women (n = 40).
32 We examined changes from pre- to postpartum in both objective memory performance as
33 well as subjective memory complaints. We also examined whether changes from pre-
34 conception to postpartum in encoding-related brain activity were associated with shifts in

1 objective memory performance and self-reported memory complaints. Additionally,
2 behavioral data collected during the pregnancy session were used to examine whether
3 potential changes in objective memory performance and subjective memory complaints were
4 already present during pregnancy itself. Analyses involving the late postpartum session were
5 conducted to examine the persistence of any effects across the first year of motherhood. We
6 also used a separate cohort to see if we could replicate the findings in memory performance
7 and subjective memory complaints in first-time mothers (n = 72) and nulliparous controls (n
8 = 72).

9
10 As changes in sleep, stress, and depressive symptoms are known to independently impair
11 cognitive performance during the peripartum period^{8,19,20} and may differ between first- and
12 second-time mothers, we included pre-to-postpartum changes in sleep, stress, and depression
13 scores as covariates in our primary analyses to examine whether any observed effects on
14 memory-related brain activity and behavioral performance could be attributed to pregnancy
15 itself rather than these contextual factors. Specifically, we tested whether apparent group
16 differences, both in subjective memory complaints and in memory-related brain activity, are
17 better accounted for by these contextual, wellbeing-related factors than by pregnancy or
18 parity itself. This study provides the first longitudinal characterization of task-evoked
19 memory-related brain function across the transition to motherhood, offering a neurobiological
20 perspective on the notion of 'mommy brain'.

Method

Participants and Study Design

We employed a longitudinal design with three groups of female participants: first-time mothers (PRG₁, n = 40), second-time mothers (PRG₂, n = 30), and nulliparous controls (CTR, n = 40). Participants in the PRG₁ and PRG₂ group visited at four timepoints: before pregnancy (PRE), during pregnancy (week 36) (PRG), early postpartum (POST), and late postpartum (POST+1). CTR participants visited at similar intervals for the PRE, PRG, and POST session. As the ethics committee did not approve the inclusion of MRI scanning during pregnancy, no MRI data were acquired during the PRG session. One participant from PRG₂ did not complete the retrieval task outside the scanner at the POST session, so these data were excluded from the forgotten>remember contrast (see below). PRG₁ (n = 29) and PRG₂ (n = 14) participants also visited during the late postpartum period (POST+1). As only 14 women could be included in the PRG₂ group for the POST+1 session due to COVID-19 restrictions, we excluded this timepoint for this group from the analyses. For an overview of the timeline of the study, see Figure 1A. Other measures of this cohort (structural, magnetic resonance spectroscopy, diffusion-weighted and resting-state) have been reported in previous research^{2,3}. For more information on the inclusion, see Supplementary Figure 1.

There were no statistically significant differences in age ($p = 0.98$, $t = 0.032$), level of education ($p = 0.14$, $t = -1.51$), and PRE-to-POST time interval ($p = 0.072$, $t = 1.83$) between PRG₁ and CTR. In addition, between PRG₁ and PRG₂, there was no significant differences in level of education ($p = 0.53$, $W = 647.5$) and PRE-to-POST time interval ($p = 0.86$, $W = 585$). Between PRG₂ and CTR, there were also no differences in level of education ($p = 0.076$, $W = 730$) and PRE-to-POST time interval ($p = 0.25$, $W = 502.5$). As expected, a significant difference was observed in age between PRG₁ and PRG₂ ($p < 0.001$, $W = 930$). In addition, a significant age difference was found between PRG₂ and CTR ($p < 0.001$, $W = 930$). Hence, age was included as a nuisance covariate in all analyses. For means and standard deviations of age, level of education, time between PRE-to-POST, time between parturition-POST, time between parturition-POST+1, medical history, conception method, type of delivery, delivery of twins, sex of the child, and breastfeeding, see Supplementary Table 1.

To replicate the findings on memory performance and subjective memory complaints, we additionally drew on data from an independent longitudinal cohort collected in our group comprising first-time mothers (PRG₁; $n = 87$; mean age = 30.9 years, $SD = 3.1$) and nulliparous controls (CTR; $n = 79$; mean age = 27.8 years, $SD = 7.0$). Groups differed significantly in age ($t(106.1) = 3.70$, $p < 0.001$), and age was therefore included as a covariate in all replication analyses. This cohort followed the same longitudinal design and timepoints (PRE, PRG, POST) and applied the same inclusion criteria as the main sample. Ethical approval was granted by the METC Amsterdam UMC. There was no participant overlap with the main sample. The set-up was similar to the main study, except that the memory performance task was administered outside the scanner at each timepoint. A total of 144 participants visited the PRE and POST session (PRG₁, $n = 72$; CTR, $n = 72$), and 124 for the PRE and PRG session (PRG₁, $n = 77$; CTR, $n = 47$).

Memory Task Paradigm

Encoding Phase

During MRI scanning, participants completed an associative memory encoding task based on the Wechsler Verbal Paired Associates paradigm³¹. The task presented 70 unrelated word pairs, each displayed for 3 seconds. Stimuli consisted of 400 (100 per session) high-frequency Dutch nouns (4-7 letters) selected from the SUBTLEX_NL database³². Following this database selection, three researchers independently reviewed the candidate word list to exclude items judged to be emotionally salient, highly abstract, or otherwise idiosyncratic in meaning. During each session, participants were presented with new word pairs, resulting in four parallel versions of the task, which were matched for frequency and word length and counterbalanced across participants to prevent learning effects. A control task was also used that comprised the presentation of words where participants were asked to count the number of letters on the screen.

The task employed an event-related design with optimized null trials which were presented as a black screen (optseq; www.nitrc.org/projects/optseq). Each trial began with a 500 ms cue ("remember" or "count") followed by a 200 ms interval before presenting two words for 3000 ms. Participants then responded to either "will you remember this?" or "did you see more or less than 10 letters?" depending on the condition, for 2000 ms, ensuring task engagement and attention. A schematic representation of the experimental paradigm is provided in Figure 1B.

As we were not able to scan during the PRG session, participants performed the encoding phase outside the scanner.

In the replication cohort, the encoding phase was conducted outside the scanner at all timepoints and we used 40 word pairs instead of 70 word pairs.

Retrieval Phase

Approximately 10 minutes post-encoding, participants completed three sequential memory tests outside the MRI scanner. The first test assessed item recognition, wherein participants identified 70 previously seen words (one from each encoded pair) interspersed with 30 novel words. The second test involved cued recall, in which participants were presented with one word from each pair and asked to type its associated word. The final test examined paired recognition, where participants were shown one word from each pair alongside two other optional words and asked to select the correct paired associate. A schematic representation of these memory tests is provided in Figure 1B.

Memory Score Calculation

To enable balanced fMRI analyses despite performance variability across different retrieval tasks, we developed a composite memory score for each encoded word pair, adapting an approach similar to that described by van Kesteren et al. (2020)³³. Specifically, some participants performed at ceiling or floor on individual retrieval tasks (i.e., correctly retrieving all or none of the items), which would have rendered single-task parametric modulator analyses uninformative for those individuals. To address this, each trial received a composite score from 0 to 3, reflecting cumulative performance across the three retrieval tasks (0 = failed all tasks; 3 = succeeded in all tasks).

Subjective Memory Assessment

Subjective memory complaints were assessed using the language and verbal memory subscales of the Multiple Ability Self-report Questionnaire³⁴(MASQ). Participants rated perceived difficulties on a 5-point Likert scale ("never" to "always") for items such as "I am able to recall the details of the evening news report several hours later." Higher scores indicated greater self-reported memory deficits.

Sleep, stress, depression

We included measures of sleep quality, distress level, and depressive scores as covariates as

they have been shown to impact cognitive performance both outside the maternal context as well as during the peripartum¹⁹. Sleep was measured as the average hours of sleep in the past week, psychological distress using the 10-item Kessler Psychological Distress (K10) Scale³⁵, and depression using the Edinburgh Postpartum Depression Scale (EPDS)³⁶. In the replication cohort, we did not assess stress with the K10, but with the Perceived Stress Scale³⁷.

Neuroimaging Acquisition and Analysis

MRI acquisitions were obtained on a 3-Tesla Philips MRI scanner. High-resolution 3D T1 weighted images in transverse orientation (Repetition Time (TR) = 9.8 ms; Echo Time (TE) = 4.6 ms; Flip Angle = 8°; voxel size = 0.875 x 0.875 x 1.20 mm³; Field of View = 178 x 224 x 168 mm) and T2*-weighted whole-brain echoplanar images (EPI) were acquired (TR = 2.2s; TE = 30 ms; Flip Angle = 80°; Field of View = 220 x 220 x 111.65; voxel size = 2.75 x 2.75; 37 descending slices: 398 volumes). MRI scans were visually checked for quality control before data processing, but no images had to be excluded.

Preprocessing

Preprocessing was performed using the Data Processing Assistant for Resting-State fMRI (<http://rfmri.org/DPARSF>)³⁸, which is based on Statistical Parametric Mapping 12 (SPM 12; <http://www.fil.ion.ucl.ac.uk/spm>) in MATLAB 2022a, and the toolbox for Data Processing & Analysis of Brain Imaging, advanced edition version 5.0 (<http://rfmri.org/DPABI>)³⁹. Time series of images for each participant were realigned to the mean functional image. To remove nuisance signals, the Friston 24-parameter model was utilized to regress out head motion effects from the realigned data. Normalization was executed by the DARTEL tool⁴⁰ to transform the functional data from individual native space to MNI space (3 x 3 x 3 mm³). Spatial smoothing with a FWHM kernel of 8 mm was applied to the functional images. First-level analyses included spike regressors for head movements exceeding 3 mm.

fMRI Analyses

First-level general linear models were specified for each participant separately. The primary model included two onset regressors: one for encoding trials (word pair encoding) and one for counting trials (character counting), each convolved with the canonical hemodynamic response function (HRF). The encoding > counting contrast was used to examine overall encoding-related brain activity relative to the control task.

To examine neural correlates of successful memory encoding, a parametric modulation approach was adopted following Soch et al. (2021)³⁰, which has been shown to outperform categorical subsequent memory models (remembered > forgotten) through cross-validated Bayesian model selection. In the primary model, a parametric modulator reflecting the composite memory score (ranging from 0 to 3, reflecting performance across paired recognition, item recognition, and cued recall, see above) was added to the encoding regressor. Following Soch et al. (2021)³⁰, this score was arcsine-transformed prior to inclusion as a parametric modulator, assigning proportionally greater weight to trials with strong memory outcomes (score = 3) relative to moderate outcomes (scores 1-2), consistent with evidence that high-confidence memory signals engage memory-related brain regions more robustly. Orthogonalization of the parametric modulator with respect to the onset regressor was disabled (orth = 0) in SPM, following recommendations by Mumford et al. (2015)⁴¹ and Soch et al. (2021)³⁰, to prevent the removal of shared variance between the onset regressor and the parametric modulator and to ensure unbiased parameter estimates across participants.

A second-level model only including the PRE session was used to examine 1. whether the brain activity in all groups was in line with brain activity found in similar previous studies and 2. whether there were differences between groups already at baseline. We included age, sleep, stress, and depression scores at the PRE session as covariates in the model.

Then, to answer our main question, which was to examine whether there are differences between groups in the change in encoding-related brain activity before and after pregnancy (or at similar timepoints for the control group) for the second-level analyses, we included all three groups to examine differences in change in brain function from the PRE to the POST session. We included age at baseline as well as PRE-to-POST changes in sleep, stress, and depression as covariates in the model. We also ran the analysis without sleep, stress, and depression as covariates.

For all models, we used a voxel-level family-wise error (FWE) correction ($p < 0.05$) in line with Soch et al. 2021³⁰. Brain regions were labeled using the Automated Anatomical Labeling (AAL) toolbox with local maxima labeling and the ROI_MNI_V7 atlas template⁴².

Association between brain function and subjective and objective performance

We extracted beta-weights using a meta-analytic mask from Kim et al. 2024⁴³ that includes

the left inferior frontal and angular gyrus. We applied the mask on the main model (i.e. total memory score including sleep, stress, and depression as covariates) and extracted beta-values with MarsBar (<http://marsbar.sourceforge.net/>)⁴⁴ for all three groups. In RStudio (Version 1.4.17), we then used Pearson correlations in R to examine associations between these extracted beta-weights and both objective memory performance (cued recall, item recognition, and paired recognition) and subjective memory complaints (MASQ language and verbal memory subscales). To correct for multiple comparisons across the five outcome measures, we applied FDR correction to the resulting p-values.

To evaluate whether the null result for this primary Group $\times$ Session brain contrast reflected a genuine absence of an effect rather than insufficient power, given that whole-brain voxelwise equivalence testing lacks an established multiple-comparisons framework, we applied a sensitivity analysis, equivalence testing, and Bayesian model comparison to the region-of-interest beta-weights described above, reflecting the PRE-to-POST change in Kim et al. (2024) mask activity for each participant. Sensitivity analyses determined the smallest effect size (Cohen's f) each model had 80% power to detect at the relevant Bonferroni-corrected alpha, given the achieved sample size (R package *pwr*). Equivalence testing was performed on the partial eta squared of the Group $\times$ Session interaction using `TOSTER::equ_anova()`, following the approach of Campbell and Lakens (2021)⁴⁵, with the equivalence bound set to the effect size identified by the sensitivity analysis, in line with the resource-based approach recommended by Lakens (2017). As a complementary Bayesian approach, we computed Bayes factors comparing models with and without the Group $\times$ Session interaction term using the R package *BayesFactor* (default JZS prior, medium width), with participant included as a random factor; these Bayesian models did not include the covariates used in the frequentist ANCOVAs.

Statistical analyses of subjective and objective memory performance

To examine behavioral changes across the same timepoints as the neuroimaging analyses, we conducted mixed repeated-measures ANCOVAs with group (PRG₁, PRG₂, CTR) as the between-subjects factor and Session (PRE, POST) as the within-subject factor. Age at baseline and changes in psychological distress, sleep, and depression (POST minus PRE scores) were included as covariates. These analyses were conducted for objective memory performance (cued recall, item recognition, and paired recognition) and subjective memory

complaints (MASQ Verbal Memory and Language subscales) separately in RStudio (Version 1.4.17).

To examine whether any changes were already present during pregnancy, prior to the postpartum sessions included in the primary analyses, we repeated this approach using only the PRE and PRG timepoints (with PRG minus PRE covariate difference scores). Note that these data were acquired outside of the MRI scanner, in contrast to the PRE and POST sessions.

To examine whether postpartum changes persisted into the late postpartum period, we additionally analyzed the PRE, POST, and POST+1 timepoints. As sufficient data for the POST+1 timepoint were only available for the PRG₁ group, this analysis was restricted to PRG₁ using one-way repeated measures ANOVAs.

To correct for multiple testing across the three memory performance outcomes and two subjective memory outcomes respectively, we applied Bonferroni correction, which controls the familywise error rate and is well-suited to a small number of theoretically motivated primary comparisons. This resulted in a corrected alpha of .017 for the three objective memory outcomes and 0.025 for the two subjective memory outcomes. Significant interaction effects were followed up with Tukey's HSD post-hoc analyses. Effect sizes are reported as partial eta squared (η^2_p).

To examine whether participants showed consistent patterns of memory change across the three retrieval components, we computed Pearson correlations between the PRE-to-POST difference scores for cued recall, item recognition, and paired recognition for all groups. Positive correlations would indicate that individuals who improved (or declined) on one memory component tended to do so similarly across the other components, reflecting a common underlying change rather than component-specific variation.

Subjective-objective memory associations

To examine associations between subjective memory complaints and objective memory performance across the peripartum transition, PRE-to-POST difference scores were computed for each measure separately. Objective memory performance was assessed across three tasks: cued recall, item recognition, and paired recognition, with higher scores indicating better performance. Subjective memory complaints were assessed using two MASQ subscales:

1 language and verbal memory, with higher scores indicating greater complaints. For each
2 group (PRG₁, PRG₂, CTR), Pearson or Spearman correlations were computed between each
3 objective and each subjective difference score, yielding six correlations per group (18 in
4 total). Prior to each correlation, normality of both variables was assessed using Shapiro-Wilk
5 tests; Spearman rank correlation was applied if either variable violated normality ($p < 0.05$).
6 To control for multiple comparisons, p-values were corrected using the Benjamini-Hochberg
7 false discovery rate (FDR) procedure across all 18 tests.

8 **Statistical analyses of psychological distress, sleep, and depression**

9 To examine whether there were changes in psychological distress, sleep, and depression, we
10 used repeated measures ANOVA with group (PRG₁, PRG₂, CTR) as the between-subjects
11 factor and timepoint (PRE, POST) as the within-subject factor for each measure. Significant
12 interaction effects were followed up with Tukey's HSD post-hoc analyses.

13

Results

fMRI results

Significant activation during successful memory encoding was observed in the inferior frontal gyrus (orbital part), middle temporal gyrus, and parahippocampal gyrus, consistent with previous meta-analytic findings on successful verbal memory encoding⁴³, and in the precuneus, posterior cingulate cortex, and supramarginal gyrus, regions previously implicated in episodic memory encoding⁴⁶. We also observed activation in frontal regions, the caudate nucleus, and the middle occipital gyrus. See Table 1 and Figure 2 for an overview. We observed no significant differences at baseline between groups in brain activation during successful memory encoding.

Furthermore, we tested whether the transition of motherhood was associated with differential neural changes in verbal memory encoding by comparing the changes in neural activity between our participant groups. These analyses involved a parametric modulation of the total memory score during encoding relative to the counting control task, including covariates age, and changes in sleep, psychological stress, and depression scores. When examining the changes from pre- to post-pregnancy in the first-time and second-time mothers, and at similar timepoints for control women, we observed no significant differences between any groups. In the model without sleep, psychological stress, and depression scores added as covariates, there were also not significant group differences.

To further evaluate this null result, we conducted a sensitivity analysis, equivalence testing, and Bayesian model comparison on the region-of-interest beta-weights reflecting the PRE-to-POST change in Kim et al. (2024) mask activity. The design had approximately 80% power to detect a group effect of $\eta^2_p \approx 0.10$. The observed group effect was small ($F(2, 85) = 0.16$, $p = 0.849$, $\eta^2_p = 0.004$) and was statistically equivalent to zero against this bound ($p_{eq} = 0.0015$). Bayesian model comparison (BayesFactor, default JZS prior) yielded a Bayes factor of $BF_{01} \approx 8.65$ in favor of the model without a group effect, providing moderate evidence for the absence of a group difference in this primary region of interest.

Association between brain function and subjective and objective performance

We extracted beta-weights from our main brain model examining the changes from PRE-to-POST in successful memory encoding, with a meta-analytic mask by Kim et al. 2024⁴³ applied and correlated these with changes in memory performance and self-reported memory

changes. Beta-weights were positively correlated with post-minus-pre change scores in item recognition in first-time mothers ($r = 0.46$, $t(28) = 2.77$, $p = 0.010$, 95% CI [0.12, 0.71], $p_{FDR} = 0.029$). A trend was observed for cued recall ($p = 0.39$, $p = 0.033$), which did not survive FDR correction ($p_{FDR} = 0.050$). No significant correlations were observed in PRG₂ or CTR, see Supplementary Table 2. In addition, no significant correlations were found between beta-weights and subjective memory scores for any group, see Supplementary Table 3. Because this correlation was significant only in PRG₁, we directly tested whether it differed from the correlations observed in the other two groups using Fisher r-to-z comparisons. The correlation in PRG₁ differed from that in PRG₂ at an uncorrected threshold ($z = 2.07$, $p = .038$) but not after Bonferroni correction for the three pairwise comparisons ($p = .114$), and did not differ significantly from CTR (uncorrected $z = 1.68$, $p = .094$; PRG₂ versus CTR: $z = -0.50$, $p = .615$). Given this, together with the fact that this correlation was one of several tested, we interpret this finding as an exploratory, descriptive observation rather than evidence for a PRG₁-specific effect.

Behavioral changes across pregnancy

For means and standard deviation for both memory performance and subjective memory complaints for each timepoint, see Supplementary Table 4 and 9, and Supplementary Figures 2-14.

Memory performance

Using our Bonferroni-corrected threshold, there was no significant group $\times$ session interaction for cued recall ($F(2, 81) = 2.58$, $p = 0.082$, $\eta^2_p = 0.060$), item recognition score ($F(2, 81) = 4.02$, $p = 0.022$, $\eta^2_p = 0.090$), and paired recognition ($F(2, 81) = 0.30$, $p = 0.74$, $\eta^2_p = 0.007$), see Figure 3 for an overview and Supplementary Table 5-7 for all statistics. To examine consistency across tasks, we ran separate correlation analyses, presenting significant correlations of POST minus PRE scores between all three tasks, see Supplementary Table 8.

In the replication cohort, group $\times$ session interactions were not significant for any memory outcome across the PRE to POST interval: cued recall, $F(1, 120) = 2.33$, $p = 0.13$, $\eta^2_p = .019$; paired recognition, $F(1, 119) = 0.92$, $p = 0.34$, $\eta^2_p = 0.008$; item recognition, $F(1, 119) = 1.37$, $p = 0.24$, $\eta^2_p = 0.011$, see Figure 3.

During pregnancy

To test if differences in memory performance were present during pregnancy, we also ran an ANCOVA model with the PRE and PRG timepoints only. For the cued recall score, the group $\times$ session interaction was significant and survived the Bonferroni-corrected threshold ($F(2, 81) = 9.25, p < 0.001, \eta^2_p = 0.19$). There was a significant main effect of session ($F(1, 81) = 43.54, p < 0.001, \eta^2_p = 0.35$) but not of group ($F(2, 81) = 0.64, p = 0.53$), see Supplementary Table 10. Post-hoc analysis shows a significant increase from PRE-to-PRG for PRG₁ ($p = 0.003$) and CTR ($p < 0.001$), but not for PRG₂ ($p = 0.34$), see Supplementary Table 11. There was no significant group $\times$ session interaction for both item recognition ($p = 0.12$) as well as the paired recognition score ($p = 0.21$), see Supplementary Table 12-13.

In the replication cohort, group $\times$ session interactions were not significant for any memory outcome across the PRE to PRG interval: cued recall, $F(1, 92) = 1.02, p = .315, \eta^2_p = .011$; paired recognition, $F(1, 82) = 0.05, p = 0.83, \eta^2_p < 0.001$; item recognition, $F(1, 91) = 1.21, p = 0.275, \eta^2_p = 0.013$.

One year postpartum

To examine within-group changes from pre-conception until the late postpartum period, analyses were performed including the POST+1 session data available only for the PRG₁ group. One-way repeated measures ANOVAs were conducted for each test outcome separately, with session (PRE, POST, POST1) as the within-subject factor. There was no significant effect that survived Bonferroni corrected threshold for cued recall ($F(2, 56) = 2.15, p = 0.126, \eta^2_p = 0.071$), item recognition ($F(1.37, 38.3) = 4.09, p = 0.038, \eta^2_p = 0.128$), and paired recognition $F(2, 56) = 2.19, p = 0.121, \eta^2_p = 0.073$).

Subjective memory complaints

For the MASQ language subscale, the group $\times$ session interaction did not reach the Bonferroni-corrected threshold ($F(2, 81) = 2.94, p = 0.058, \eta^2_p = 0.068$), Supplementary Table 14. However, for the MASQ verbal memory subscale, there was a significant group $\times$ session interaction that survived the Bonferroni-corrected threshold ($F(2, 81) = 4.83, p = 0.010, \eta^2_p = 0.11$). There was a significant main effect of session ($F(1, 81) = 53.42, p < 0.001, \eta^2_p = 0.40$), but the group main effect was not significant ($p = 0.67$), see Supplementary Table 15. Post-hoc analysis shows significant increases from PRE to POST for PRG₁ ($p < 0.001$) and PRG₂ ($p < 0.001$), but not for CTR ($p = 0.093$), see Supplementary Table 16. See Figure 4 for an overview.

In a separate larger sample, group $\times$ session interactions were significant for both MASQ language ($F(1, 125) = 10.30, p = 0.002, \eta^2_p = 0.076$) and MASQ verbal memory ($F(1, 125) = 7.04, p = 0.009, \eta^2_p = 0.053$), surviving the Bonferroni-corrected threshold ($\alpha = .025$). Post-hoc analyses showed significant increases from PRE to POST for PRG₁ (MASQ language: $p < 0.001$; MASQ verbal memory: $p < 0.001$) but not for CTR (MASQ language: $p = 0.221$; MASQ verbal memory: $p = 0.78$). Change in perceived stress (Δ PSS) was a significant predictor of change in MASQ language complaints postpartum ($F(1, 125) = 4.32, p = 0.040, \eta^2_p = 0.033$), see Figure 4.

During pregnancy

For the MASQ language subscale, the group $\times$ session interaction was significant and survived the Bonferroni-corrected threshold ($F(2, 84) = 8.53, p < 0.001, \eta^2_p = 0.17$). There was a significant main effect of session ($F(1, 84) = 64.30, p < 0.001, \eta^2_p = 0.43$), but the group main effect was not significant ($p = 0.96$), see Supplementary Table 17. Post-hoc analysis shows significant increases in subjective language complaints from PRE-to-PRG for PRG₁ ($p < 0.001$) and PRG₂ ($p < 0.001$), but not for CTR ($p = 0.056$), see Supplementary Table 18. Among the covariates, change in depression score (Δ EPDS) was a significant predictor of the change in MASQ Language subscale score ($F(1, 84) = 4.15, p = 0.045, \eta^2_p = 0.047$).

For the MASQ verbal memory subscale, there was also a significant group $\times$ session interaction that survived the Bonferroni-corrected threshold ($F(2, 84) = 8.27, p < 0.001, \eta^2_p = 0.16$). There was a large significant main effect of session ($F(1, 84) = 69.89, p < 0.001, \eta^2_p = 0.45$), but the group main effect was not significant ($p = 0.77$), see Supplementary Table 19. Post-hoc analysis shows significant increases from PRE-to-PRG for PRG₁ ($p < 0.001$) and PRG₂ ($p < 0.001$), but not for CTR ($p = 0.14$), see Supplementary Table 20. Among the covariates, change in depression (Δ EPDS) was a significant predictor of the change in MASQ verbal subscale score ($F(1, 84) = 4.18, p = 0.044, \eta^2_p = 0.047$). In addition, age was also a significant predictor of the change in MASQ verbal subscale score ($F(1, 84) = 7.58, p = 0.0072$).

These findings were largely replicated in an independent sample ($n = 106$; 60 PRG₁, 46 controls) using perceived stress (Δ PSS) as an additional covariate in place of K10. For the MASQ language subscale, the group $\times$ session interaction was significant during pregnancy ($F(1, 100) = 4.63, p = 0.034, \eta^2_p = 0.044$), with pregnant participants showing a significant

increase in subjective language complaints from PRE to PRG ($p < 0.001$) while controls did not ($p = .920$). Notably, change in perceived stress (ΔPSS) was a significant predictor of change in subjective memory complaints ($F(1, 100) = 6.19, p = .015, \eta^2_p = .058$). For the MASQ verbal memory subscale, the group $\times$ session interaction did not reach significance during pregnancy ($F(1, 100) = 3.81, p = 0.054, \eta^2_p = 0.037$).

One year postpartum

Results indicated that subjective memory complaints remained elevated across the postpartum period, with no significant attenuation from early to late postpartum. For the MASQ language subscale, there was a significant effect of session that survived the Bonferroni-corrected threshold ($F(2, 46) = 7.90, p = 0.0010, \eta^2_p = 0.26$). Post-hoc analysis shows significant increases from PRE to POST ($p = 0.0040$) and from PRE to POST1 ($p = 0.016$), and no significant difference between POST and POST1 ($p = 0.91$), see Supplementary Table 21.

For the MASQ verbal memory subscale, there was also a significant effect of session that survived the Bonferroni-corrected threshold ($F(2, 46) = 7.46, p = 0.002, \eta^2_p = 0.245$). Post-hoc analysis shows significant increases from PRE to POST ($p = 0.003$) and from PRE to POST1 ($p = 0.029$), and no significant difference between POST and POST1 ($p = 0.48$), see Supplementary Table 22.

Association subjective and objective memory

Correlations between objective memory performance and subjective memory complaint difference scores are presented in Supplementary Table 23. After FDR correction, one significant negative correlation was observed: in CTR, paired recognition change scores were negatively associated with MASQ Language change scores ($r = -0.52, p = 0.002, p_{FDR} = 0.034$).

Sleep, distress, depression changes across pregnancy

For mean and standard deviations, see Supplementary Table 24-25, and Supplementary Figure 15-18. ANOVA results are presented in Supplementary Tables 26-28. Analysis of sleep patterns revealed a significant group $\times$ timepoint interaction ($p = 0.044$), with groups differing at POST ($p = 0.022$) such that CTR showed higher sleep hours than PRG₂. Notably, only the PRG₁ group demonstrated a significant decrease from pre-conception to postpartum

- 1 (p = 0.015). There was no significant group x timepoint interactions for psychological
- 2 distress ($F(2, 107) = 0.30$, $p = 0.74$) or depression scores ($F(2, 107) = 0.40$, $p = 0.68$).

Discussion

In this first longitudinal task-based fMRI study examining memory-related brain function across the transition to motherhood, we found no evidence that becoming a mother, whether for the first or second time, is associated with changes in encoding-related brain activity during a verbal memory task. Consistent with this, objective memory performance did not differ significantly between mothers and nulliparous control women. Together, these findings do not support a neurobiological basis for the popular notion of 'mommy brain', at least with respect to verbal memory encoding. Interestingly, however, both first- and second-time mothers reported significant increases in subjective memory complaints postpartum, a pattern absent in control women.

As a first step, we examined whether the expected pattern of encoding-related brain activation was present in our sample. Successful memory encoding was associated with widespread activation across medial temporal, parietal, and prefrontal regions, consistent with previous findings on verbal associative memory encoding^{29,30,43,46}. No baseline differences between groups were observed at the pre-conception session, confirming that the groups were comparable in memory-related brain function prior to the transition to motherhood.

Against the backdrop of this well-characterized task response, we then examined whether the transition to motherhood was accompanied by changes in encoding-related brain activity from pre-conception to postpartum. No significant group differences were observed when accounting for age at baseline and pre-to-postpartum changes in sleep, psychological distress, and depressive scores. Further analyses supported that this null result reflected a genuine absence of an effect rather than insufficient power: equivalence testing indicated that the observed group difference was substantially smaller than what our design had the power to detect, and a Bayesian model comparison provided further evidence in favor of no group effect over a model that included one. This suggests that neither first- nor second-time motherhood is associated with differential changes in the neural substrates of verbal memory encoding relative to nulliparous control women.

Consistent with the absence of group differences in encoding-related brain activity, no significant group $\times$ session interaction was observed for any of the three memory tasks, suggesting that the transition to motherhood does not produce reliable group-level differences

1 in verbal memory relative to nulliparous control women. These findings were replicated in a
2 larger independent sample that included women becoming first-time mothers ($n = 72$) and
3 nulliparous control women ($n = 72$). This is consistent with several previous studies: Orchard
4 et al.⁸ found no objective differences in verbal memory, working memory, or processing
5 speed between first-time mothers and non-mothers at one year postpartum, and Christensen et
6 al. (2010)¹⁸ similarly reported largely non-significant differences in memory recall, cognitive
7 speed, and working memory in the largest prospective cohort study to date. A recent
8 systematic review by Younis et al.¹¹ found only modest, heterogeneous effects on verbal
9 memory and attention during pregnancy, with considerable variability across studies.
10 Similarly, Raz (2024)⁴⁷ found no evidence of memory decline across a comprehensive battery
11 of verbal, visuospatial, and executive memory tasks in late pregnancy, and additionally
12 reported enhanced inhibitory control in pregnant women, accompanied by increased P300
13 amplitude during a stop-signal task, suggesting that pregnancy-related neural changes may be
14 domain-specific rather than reflecting general cognitive decline. Callaghan et al. (2022)¹⁰
15 reported enhanced, rather than impaired memory scores during pregnancy for a different type
16 of memory, spatial associative memory, particularly for parenting-relevant stimuli, in line
17 with an adaptationist perspective. Together, these studies illustrate the inconsistency in the
18 existing literature, which likely reflects substantial methodological variability in tasks used,
19 timing of assessment, and failure to account for confounding factors such as sleep, distress,
20 and depression by some studies. Notably, our analyses directly addressed these limitations by
21 including pre-to-postpartum changes in these factors as covariates, and no group differences
22 emerged.

23 For the session during pregnancy, cued recall showed a significant improvement for first-time
24 mothers and controls, but not for second-time mothers. However, these data were collected
25 outside the MRI scanner, in a different context from the other sessions, which may
26 independently have affected performance, for instance through reduced environmental
27 distraction. Importantly, this pattern was not replicated in our replication cohort, in which all
28 tasks were administered outside the scanner across all timepoints, showing no significant
29 changes in memory performance for any of the tasks either during pregnancy or postpartum.

30 This discrepancy illustrates a broader point: task and testing-context variability may shape
31 whether pregnancy-related changes in memory performance are detected. It remains an open
32 question whether a more ecologically valid or naturalistic memory paradigm, of the kind that

has revealed enhanced retention of parenting-relevant stimuli during pregnancy in other work, or another memory domain, would be more sensitive to change than the laboratory verbal paired-associate task used here, and whether such a paradigm might also reveal corresponding neural changes that this task-evoked fMRI approach did not detect.

In contrast to the absence of objective differences, both first- and second-time mothers showed significant increases in subjective memory complaints across the transition to motherhood, while control women did not, resulting in significant group $\times$ session interactions. The absence of a similar increase in control women, who experienced comparable follow-up intervals, argues against a simple test-retest or general aging explanation. These findings were replicated in a larger independent sample, in which the group $\times$ session interactions were significant for both MASQ subscales postpartum, and for the MASQ language subscale during pregnancy, with effect sizes that were comparable to or larger than those observed in the main sample. These findings are further consistent with Orchard et al.⁸, who found that first-time mothers reported significantly worse subjective memory than non-mothers at one year postpartum, despite showing no objective differences on measures of verbal memory, working memory, or processing speed. Increases in subjective memory complaints were observed both during pregnancy and postpartum, indicating that the elevation in subjective complaints emerges early in the transition to motherhood and persists beyond childbirth.

The pattern of covariate associations points toward wellbeing, rather than the biological changes during pregnancy itself, as a key driver of subjective memory complaints. In our main sample, change in depressive scores was a significant predictor of change in subjective memory complaints during pregnancy for both MASQ subscales. This is consistent with Orchard et al.⁸, who found that in new mothers, poorer subjective memory was associated with worse sleep, higher anxiety, and higher depression, and that the group difference in subjective memory was no longer significant after adjusting for these wellbeing factors. In the replication sample, where perceived stress was included as a covariate instead of general psychological distress that was used in the main sample, it was change in stress rather than depressive scores that emerged as a significant predictor of subjective memory complaints, both during pregnancy and postpartum. Taken together, these findings across two independent samples suggest that subjective memory complaints in the perinatal period might not be a direct reflection of memory ability, but are instead closely tied to broader indices of

psychological wellbeing. The specific wellbeing factor that drives this association may depend on the measure used, but the pattern points consistently toward mood and stress as mediating variables.

In the present study, pre-to-postpartum objective performance was also not linked with subjective memory complaints in the pregnancy groups, but changes in sleep, psychological distress, and depressive scores did not account for the group difference in subjective memory complaints, as the interaction remained significant after including these as covariates. However, age and depressive scores were associated with changes in subjective memory complaints across both MASQ subscales from pre-conception to pregnancy, suggesting that individual differences in mood and life stage contribute to early increases in the perception of memory difficulties during pregnancy. Together, these results suggest that subjective memory complaints are a genuine and specific feature of the transition to motherhood, in the absence of a corresponding decline in objective verbal memory performance. This was further corroborated by the absence of significant correlations between objective memory change scores and subjective memory complaint change scores in either maternal group, while in control women, improvements in paired recognition were associated with reductions in language-related memory complaints.

Analyses of first-time mothers at the late postpartum session further indicated that objective memory performance remained stable while subjective complaints stayed elevated one year after birth, indicating that elevated subjective complaints persist well into the postpartum period even as objective performance remains stable. This is consistent with Orchard et al. (2022)⁸, who similarly found elevated subjective memory complaints in the absence of objective deficits at one year postpartum, and who noted that the persistence of such complaints may reflect sustained increases in cognitive load and heightened awareness of memory lapses rather than true memory decline. Although the absence of a control group at the late postpartum timepoint warrants caution in interpretation, the convergence of findings across studies suggests that subjective memory complaints may be a lasting feature of early motherhood that does not resolve in the months following childbirth.

Beyond the wellbeing-related factors identified above, social psychological frameworks offer a complementary perspective on why subjective memory complaints might arise and persist in the absence of objective decline. Pownall et al. (2022)⁴⁸ argue that inconsistencies in the 'baby brain' literature may partly reflect unexamined social and contextual influences on how

pregnant and postpartum women appraise their own cognitive functioning, and propose stereotype threat theory and objectification theory as candidate explanatory frameworks. According to stereotype threat theory, exposure to the pervasive cultural stereotype that pregnancy and motherhood impair memory could itself bias self-perceived cognitive functioning, independent of any underlying neurobiological change. Objectification theory similarly suggests that the bodily scrutiny to which pregnant and postpartum women are frequently subject may deplete cognitive resources through self-objectification, again without this necessarily reflecting a true decline in memory ability.

Although no group differences in encoding-related brain activity were observed, brain activity within the meta-analytically defined memory network was associated with memory performance change at the individual level in first-time mothers. Specifically, greater increases in encoding-related brain activity from pre-conception to postpartum were associated with greater improvements in item recognition performance, and a similar trend was observed for cued recall. This suggests that individual variability in neural change during the transition to first-time motherhood is meaningfully related to memory outcomes, even in the absence of group-level effects. However, we cannot explain why we did not find comparable associations between objective memory performance and brain response in the other two groups, which warrants further investigation.

The present findings speak to a broader question in the field: whether the well-documented structural and resting-state brain changes observed across the transition to motherhood reflect cognitive decline or neural adaptation. Here, for the first time, we examine task-based brain responses during this transition, and find no evidence that encoding-related brain activity or objective memory performance differs between women becoming mothers and nulliparous controls, thus providing no support for a neurobiological decline account. It is important to note, however, that these conclusions are specific to the domain of verbal associative memory and future studies need to examine changes in brain function for other cognitive domains to examine potential cognitive decline across pregnancy and postpartum. Also, research using emotionally valenced material and parenting-relevant stimuli suggests that pregnancy-related neural and cognitive changes may be domain-specific and not fully captured by neutral laboratory tasks^{10,47}, with motivation and selective attention toward infant-related information representing additional factors that may shape memory performance in mothers. Whether the present findings extend to other domains such as social cognition, which may be more

1 directly relevant to maternal behavior, remains an important open question for future task-
2 based neuroimaging research.

3 Several methodological considerations warrant acknowledgment. It should be noted that the
4 pregnancy session was conducted outside the scanner because scanning during pregnancy
5 was not allowed by the ethical committee. This likely enhanced memory encoding due to
6 reduced environmental distractions, resulting in elevated memory scores across all groups
7 that subsequently declined at postpartum. In addition, our study sample had a relatively high
8 education level, which, although consistent across groups, may limit the generalizability of
9 findings to women with more diverse educational backgrounds. Furthermore, although mean
10 depression scores did not exceed clinical thresholds, they were relatively elevated across all
11 groups and sessions. Depression scores were included as covariates in all analyses. The sleep
12 measure, which captured hours slept in the week prior to each session, would benefit from
13 more comprehensive metrics including sleep quality and fragmentation. Future investigations
14 could also examine other contextual factors known to impact cognitive performance in
15 mothers, such as social support, mental load, and parenting demands^{8,10,19,49,50}, which
16 represent major changes in daily life that may shape memory functioning beyond what is
17 captured by mood and sleep measures alone.

18 Finally, it should be noted that our results cannot be used to dismiss the subjective memory
19 complaints experienced by pregnant women and new mothers. It is possible that, similar to
20 disorders of maternal mental health, a small subset of women experiences particularly severe
21 memory difficulties that would not be captured in a prospective cohort study of this kind. The
22 inclusion of subjective memory measures alongside objective assessment represents a
23 strength of the present study, as self-reported complaints may better reflect real-world
24 cognitive experience than laboratory task performance⁸. This is underscored by the differing
25 patterns observed here between objective and subjective memory outcomes, and highlights
26 the value of multi-method approaches in future research on maternal cognition.

27 **Conclusions**

28 This study provides the first longitudinal characterization of task-evoked brain function
29 across the transition to motherhood. Using a pre-conception design with three well-matched
30 groups followed across multiple timepoints, we found no evidence that becoming a mother,
31 whether for the first or second time, is associated with changes in verbal memory-related
32 brain activity or objective memory performance relative to nulliparous control women.

1 Notably, however, both first- and second-time mothers reported significant and persistent
2 increases in subjective memory complaints, a pattern not observed in control women. These
3 findings do not support a neurobiological basis for the popular notion of 'mommy brain' with
4 respect to verbal memory encoding.

5 **Acknowledgements**

6 We acknowledge the participants for their contribution to this study. We thank research
7 assistants A. van der Geest, A. Glasbergen, S. Altikulac, A. van Steenbergen, R. van Dort, P.
8 Berns en I. Langereis for coordinating the data collection for this project. This project was
9 supported by an Innovational Research Incentives Scheme grant (Veni, 451-14-036, E.H.) by
10 the Netherlands Organization for Scientific Research (NWO), a grant of the Leiden
11 University Fund / Elise Mathilde Fund (CWB 740s / 2t-03-2017 /EM), and a NARSAD grant
12 from the Brain and Behaviour Research Foundation, U.S.A. (grant number 25312, E.H.)
13 awarded to E. Hoekzema. E. Hoekzema is currently supported by an ERC Starting Grant
14 (948031, E.H.) provided by the European Research Council.

15 **Data availability**

16 The data used in this study, including demographic information, memory performance scores,
17 subjective memory complaints, psychological distress, sleep, and depression scores, and the
18 raw MRI images will be made available online after publication for participants who
19 provided consent for online data sharing. Preprocessed, de-identified data of the remaining
20 participants can be made available upon reasonable request.

21

1 Literature

- 2 1. Hoekzema E, Barba-Müller E, Pozzobon C, et al. Pregnancy leads to long-lasting
3 changes in human brain structure. *Nat Neurosci.* 2017;20(2):287-296.
4 doi:http://dx.doi.org/10.1038/nn.4458
- 5 2. Hoekzema E, van Steenbergen H, Straathof M, et al. Mapping the effects of pregnancy
6 on resting state brain activity, white matter microstructure, neural metabolite
7 concentrations and grey matter architecture. *Nat Commun.* 2022;13(1):6931.
8 doi:10.1038/s41467-022-33884-8
- 9 3. Straathof M, Halmans S, Pouwels PJW, Crone EA, Hoekzema E. The effects of a
10 second pregnancy on women's brain structure and function. *Nat Commun.*
11 2026;17(1):1495.
- 12 4. Pritschet L, Taylor CM, Cossio D, et al. Neuroanatomical changes observed over the
13 course of a human pregnancy. *Nat Neurosci.* Published online 2024:1-8.
14 doi:10.1038/s41593-024-01741-0
- 15 5. Servin-Barthet C, Martínez-García M, Paternina-Die M, et al. Pregnancy entails a U-
16 shaped trajectory in human brain structure linked to hormones and maternal
17 attachment. *Nat Commun.* 2025;16(1):730.
- 18 6. Paternina-Die M, Martínez-García M, Martín de Blas D, et al. Women's
19 neuroplasticity during gestation, childbirth and postpartum. *Nat Neurosci.*
20 2024;27(2):319-327. doi:10.1038/s41593-023-01513-2
- 21 7. Brown E, Schaffir J. "Pregnancy Brain": A Review of Cognitive Changes in
22 Pregnancy and Postpartum. *Obstet Gynecol Surv.* 2019;74(3):178-185.
23 doi:10.1097/OGX.0000000000000655
- 24 8. Orchard ER, Ward PGD, Egan GF, Jamadar SD. Evidence of Subjective, But Not
25 Objective, Cognitive Deficit in New Mothers at 1-Year Postpartum. *J Women's Heal.*
26 2022;31(8):1087-1096. doi:10.1089/jwh.2021.0441
- 27 9. Sharp K, Brindle PM, Brown MW, Turner GM. Memory loss during pregnancy. *BJOG*
28 *An Int J Obstet Gynaecol.* 1993;100(3):209-215. doi:10.1111/j.1471-
29 0528.1993.tb15232.x
- 30 10. Callaghan B, McCormack C, Tottenham N, Monk C. Evidence for cognitive plasticity
31 during pregnancy via enhanced learning and memory. *Memory.* 2022;30(5):519-536.

doi:10.1080/09658211.2021.2019280

11. Younis J, Bleibel M, Masri J El, Ismail A, Abou-Abbas L. Exploring the influence of pregnancy on cognitive function in women: a systematic review. *BMC Pregnancy Childbirth*. 2025;25(1). doi:10.1186/s12884-025-07181-3
12. Henry JD, Rendell PG. A review of the impact of pregnancy on memory function. *J Clin Exp Neuropsychol*. 2007;29(8):793-803. doi:10.1080/13803390701612209
13. Davies SJ, Lum JA, Skouteris H, Byrne LK, Hayden MJ. Cognitive impairment during pregnancy: a meta-analysis. *Med J Aust*. 2018;208(1):35-40. doi:10.5694/mja17.00131
14. Henry JF, Sherwin BB. Hormones and cognitive functioning during late pregnancy and postpartum: a longitudinal study. *Behav Neurosci*. 2012;126(1):73.
15. Glynn LM. Giving birth to a new brain: hormone exposures of pregnancy influence human memory. *Psychoneuroendocrinology*. 2010;35(8):1148-1155.
16. Mickes L, Wixted JT, Shapiro A, Scarff JM. The effects of pregnancy on memory: Recall is worse but recognition is not. *J Clin Exp Neuropsychol*. 2009;31(6):754-761. doi:10.1080/13803390802488111
17. Barda G, Mizrachi Y, Borokchovich I, Yair L, Kertesz DP, Dabby R. The effect of pregnancy on maternal cognition. *Sci Rep*. 2021;11(1):1-6. doi:10.1038/s41598-021-91504-9
18. Christensen H, Leach LS, Mackinnon A. Cognition in pregnancy and motherhood: prospective cohort study. *Br J Psychiatry*. 2010;196(2):126-132.
19. Orchard ER, Rutherford HJ. Birthing a new maternal cognition literature: 10 recommendations for future research. *Arch Womens Ment Health*. 2025;28(1):77-80.
20. Deater-Deckard K, Chary M, McQuillan ME, Staples AD, Bates JE. Mothers' sleep deficits and cognitive performance: Moderation by stress and age. *PLoS One*. 2021;16(1 January):1-18. doi:10.1371/journal.pone.0241188
21. Hurt NE. Legitimizing "baby brain": Tracing a rhetoric of significance through science and the mass media. *Commun Crit Stud*. 2011;8(4):376-398.
22. Crawley R, Grant S, Hinshaw KIM. Cognitive changes in pregnancy: mild decline or societal stereotype? *Appl Cogn Psychol Off J Soc Appl Res Mem Cogn*. 2008;22(8):1142-1162.

- 1 23. Pearson RM, Lightman SL, Evans J. Emotional sensitivity for motherhood: late
2 pregnancy is associated with enhanced accuracy to encode emotional faces. *Horm*
3 *Behav.* 2009;56(5):557-563.
- 4 24. Minami F, Iwata R, Nishida H, et al. Pregnancy enhances facial recognition of anger:
5 Transition from early to late pregnancy. *Eur Psychiatry.* 2024;67(S1):S809-S809.
- 6 25. Spalek K, Straathof M, Koyuncu L, et al. Pregnancy renders anatomical changes in
7 hypothalamic substructures of the human brain that relate to aspects of maternal
8 behavior. *Psychoneuroendocrinology.* 2024;164:107021.
9 doi:10.1016/j.psyneuen.2024.107021
- 10 26. Dufford AJ, Erhart A, Kim P. Maternal brain resting-state connectivity in the
11 postpartum period. *J Neuroendocrinol.* 2019;31(9):e12737.
12 doi:https://doi.org/10.1111/jne.12737
- 13 27. Rutherford HJ V, Maupin AN, Mayes LC. Parity and neural responses to social and
14 non-social stimuli in pregnancy. *Soc Neurosci.* 2019;14(5):545-548.
- 15 28. Anderson M V., Rutherford MD. Cognitive reorganization during pregnancy and the
16 postpartum period: An evolutionary perspective. *Evol Psychol.* 2012;10(4):659-687.
17 doi:10.1177/147470491201000402
- 18 29. Kim H. Neural activity that predicts subsequent memory and forgetting: A meta-
19 analysis of 74 fMRI studies. *Neuroimage.* 2011;54(3):2446-2461.
20 doi:10.1016/j.neuroimage.2010.09.045
- 21 30. Soch J, Richter A, Schütze H, et al. Bayesian model selection favors parametric over
22 categorical fMRI subsequent memory models in young and older adults. *Neuroimage.*
23 2021;230(November 2020):117820. doi:10.1016/j.neuroimage.2021.117820
- 24 31. Wechsler D. *WMS-IV: Wechsler Memory Scale.* PsychCorp; 2009.
- 25 32. Keuleers E, Brysbaert M, New B. SUBTLEX-NL: A new measure for Dutch word
26 frequency based on film subtitles. *Behav Res Methods.* 2010;42(3):643-650.
- 27 33. van Kesteren MTR, Rignanes P, Gianferrara PG, Krabbendam L, Meeter M.
28 Congruency and reactivation aid memory integration through reinstatement of prior
29 knowledge. *Sci Rep.* 2020;10(1):1-13. doi:10.1038/s41598-020-61737-1
- 30 34. Seidenberg M, Haltiner A, Taylor MA, Hermann BB, Wyler A. Development and
31 validation of a multiple ability self-report questionnaire. *J Clin Exp Neuropsychol.*

- 1994;16(1):93-104. doi:10.1080/01688639408402620
35. Kessler RC, Andrews G, Colpe LJ, et al. Kessler psychological distress scale. *Arch Gen Psychiatry*. Published online 1997.
36. Cox JL, Holden JM, Sagovsky R. Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal Depression scale. *Br J Psychiatry*. 1987;150. doi:10.1192/bjp.150.6.782
37. Cohen, S., Kamarck, T., Mermelstein R. A global measure of perceived stress. *Journal of Health and Social Behavior. J Health Soc Behav*. 1983;24(4).
38. Yan C, Zang Y. DPARSF: a MATLAB toolbox for “pipeline” data analysis of resting-state fMRI. *Front Syst Neurosci*. 2010;4. <https://www.frontiersin.org/articles/10.3389/fnsys.2010.00013>
39. Yan C, Wang X, Zuo X, Zang Y. DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics*. 2016;14(3):339-351. doi:10.1007/s12021-016-9299-4
40. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage*. 2007;38(1):95-113. doi:10.1016/j.neuroimage.2007.07.007
41. Mumford JA, Poline JB, Poldrack RA. Orthogonalization of regressors in fMRI models. *PLoS One*. 2015;10(4):1-11. doi:10.1371/journal.pone.0126255
42. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15(1):273-289.
43. Kim H. A default mode network subsystem supports both item and associative word encoding: Insights from a meta-analysis. *Imaging Neurosci*. 2024;2:imag-2.
44. Brett M, Anton J-L, Valabregue R, Poline J-B. Region of interest analysis using an SPM toolbox. In: *8th International Conference on Functional Mapping of the Human Brain*. Vol 16. Sendai, Japan; 2002:497.
45. Campbell H, Lakens D. Can we disregard the whole model? Omnibus non-inferiority testing for R² in multi-variable linear regression and in ANOVA. *Br J Math Stat Psychol*. 2021;74(1):e12201. doi:<https://doi.org/10.1111/bmsp.12201>
46. Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci*. 2008;9(8):613-625.

- 1 47. Raz S. Comprehensive assessment of memory function , inhibitory control , neural
2 activity , and cortisol levels in late pregnancy. Published online 2024:42-61.
3 doi:10.1111/nyas.15142
- 4 48. Pownall M, Conner M, Hutter RRC. ‘ Baby brain ’ in pregnancy : A review of social
5 psychological explanations and future research directions. 2022;(May):1-14.
6 doi:10.1111/spc3.12696
- 7 49. Logan DM, Hill KR, Jones R, Holt-Lunstad J, Larson MJ. How do memory and
8 attention change with pregnancy and childbirth? A controlled longitudinal examination
9 of neuropsychological functioning in pregnant and postpartum women. *J Clin Exp*
10 *Neuropsychol.* 2014;36(5):528-539. doi:10.1080/13803395.2014.912614
- 11 50. Callaghan BL, McCormack C, Kim P, Pawluski JL. Understanding the maternal brain
12 in the context of the mental load of motherhood. *Nat Ment Heal.* 2024;2(7):764-772.

13

Figure 1. A. Timeline of the included participants, timepoints, and measurements. First- and second-time mothers visited before (PRE), during (PRG), early (POST) and late postpartum (POST+1). Nulliparous controls visited at similar timepoints (PRE, PRG, POST). B. Set-up memory task. For the counting task, the instruction was: ‘count’, and the question was, ‘did you see more or fewer than 10 letters?’. The replication cohort consisted of first-time mothers and nulliparous women that visited the PRE, PRG, and POST timepoints. They conducted the entire memory task outside the scanner, and filled out questionnaires at each timepoint.

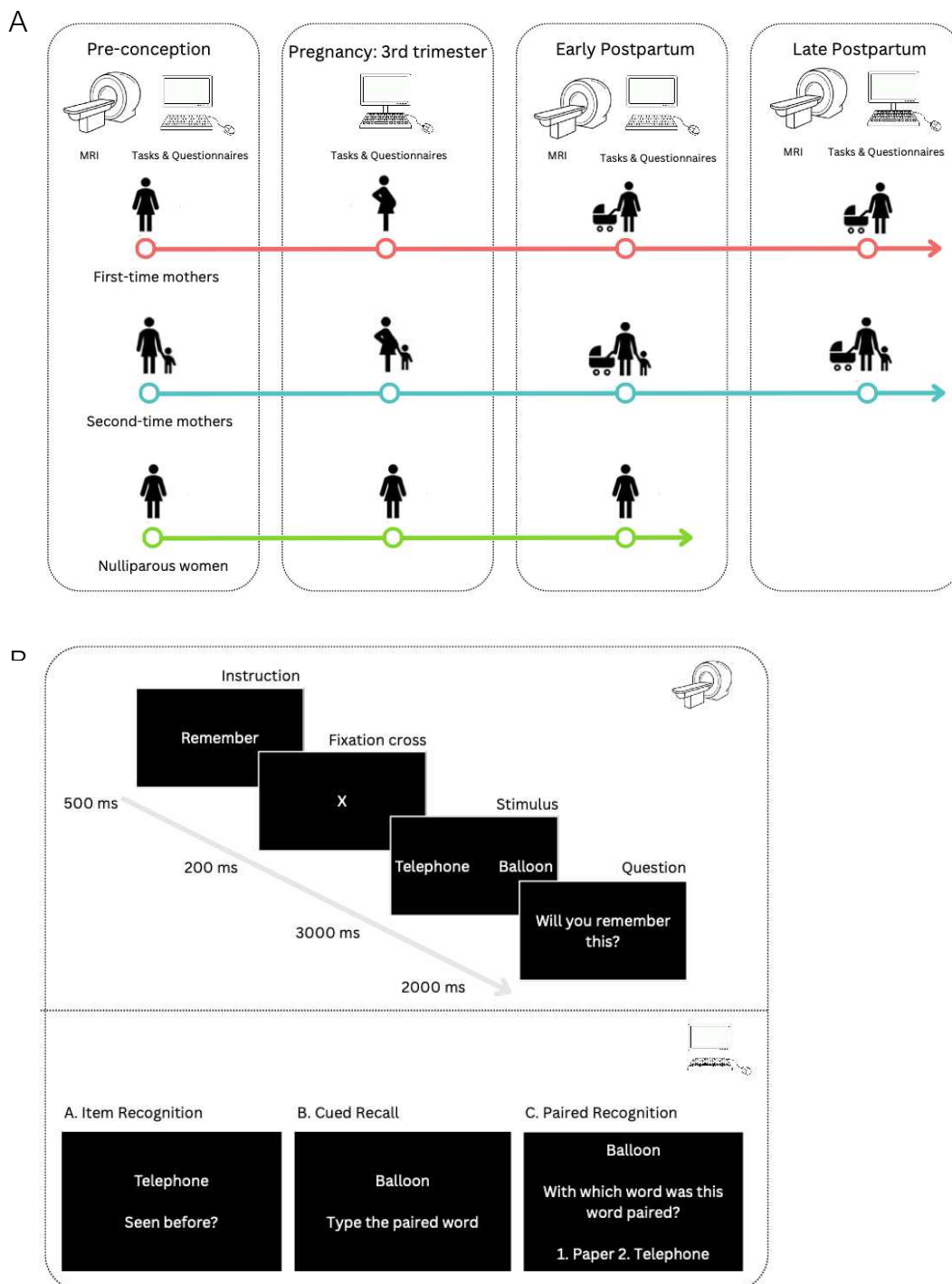


Figure 2. Whole-brain activation map depicting the average effect of condition during successful associative memory encoding, thresholded at FWE-corrected $p < 0.05$. T-values are displayed on a standard MNI152 brain using a glass brain projection (left lateral, axial, right lateral, and sagittal views).

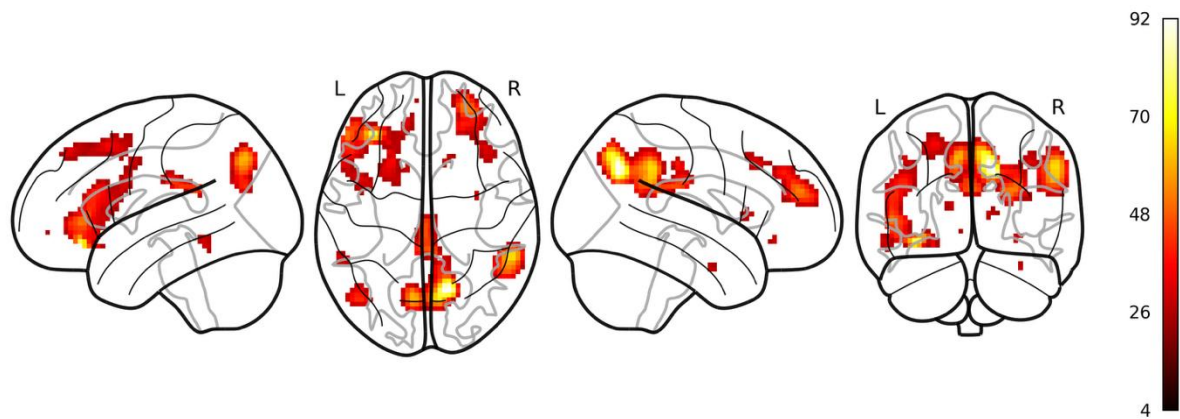


Figure 3. Objective memory performance across the transition to motherhood. Mean ($\pm$ SD) scores for Cued Recall, Item Recognition, and Paired Recognition at PRE and POST for the main cohort (Panel A; PRG₁, n = 40; PRG₂, n = 30, CTR, n = 40) and the replication cohort (Panel B; PRG₁, n = 72; CTR, n = 72). Vertical lines represent one standard deviation above and below the group mean. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls (green). PRE = before pregnancy; POST = early postpartum.

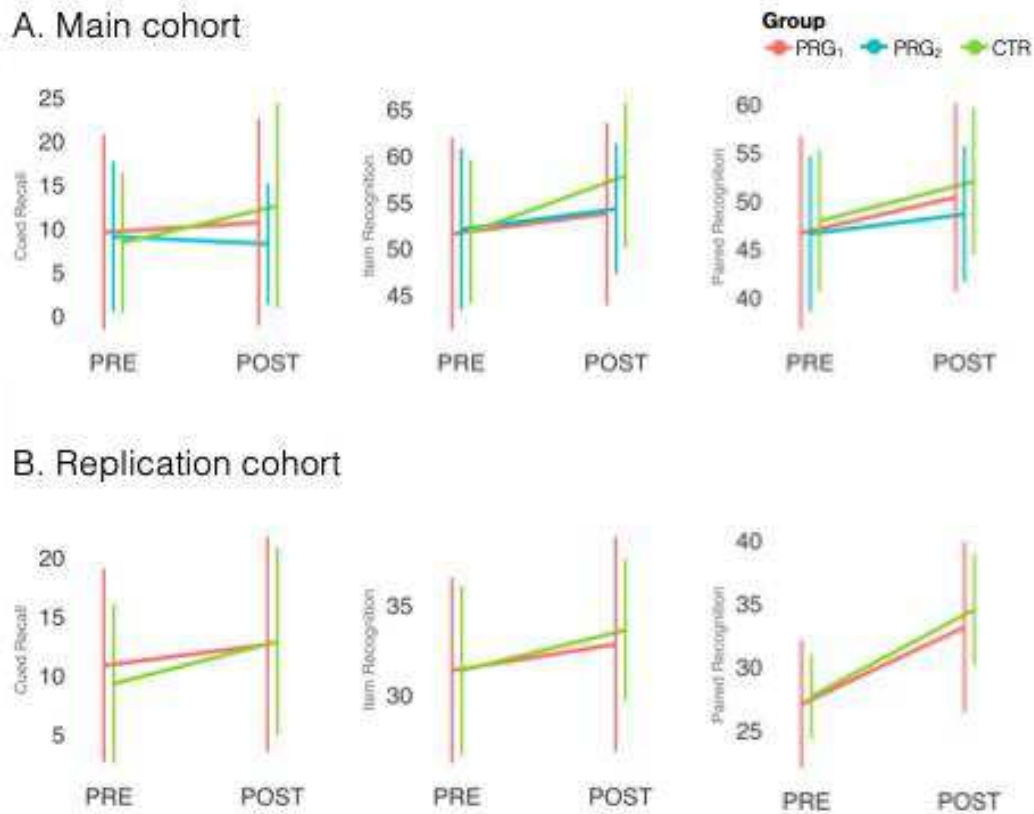


Figure 4. Subjective memory complaints across the transition to motherhood. Mean ($\pm$ SD) scores and individual trajectories for the MASQ Language subscale (Panels A and B) and MASQ Verbal Memory subscale (Panels C and D) at PRE and POST for the main cohort (Panels A and C; PRG₁, $n = 40$; PRG₂, $n = 30$, CTR, $n = 40$) and the replication cohort (Panels B and D; PRG₁, $n = 68$; CTR, $n = 64$). Left subpanels show group means $\pm$ SD; right subpanels show spaghetti plots with half-violin distributions overlaid, with individual participant trajectories in grey and the group mean trajectory in colour. Asterisks indicate significant within-group change from PRE to POST (Bonferroni-corrected, $p < .025$). PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls.

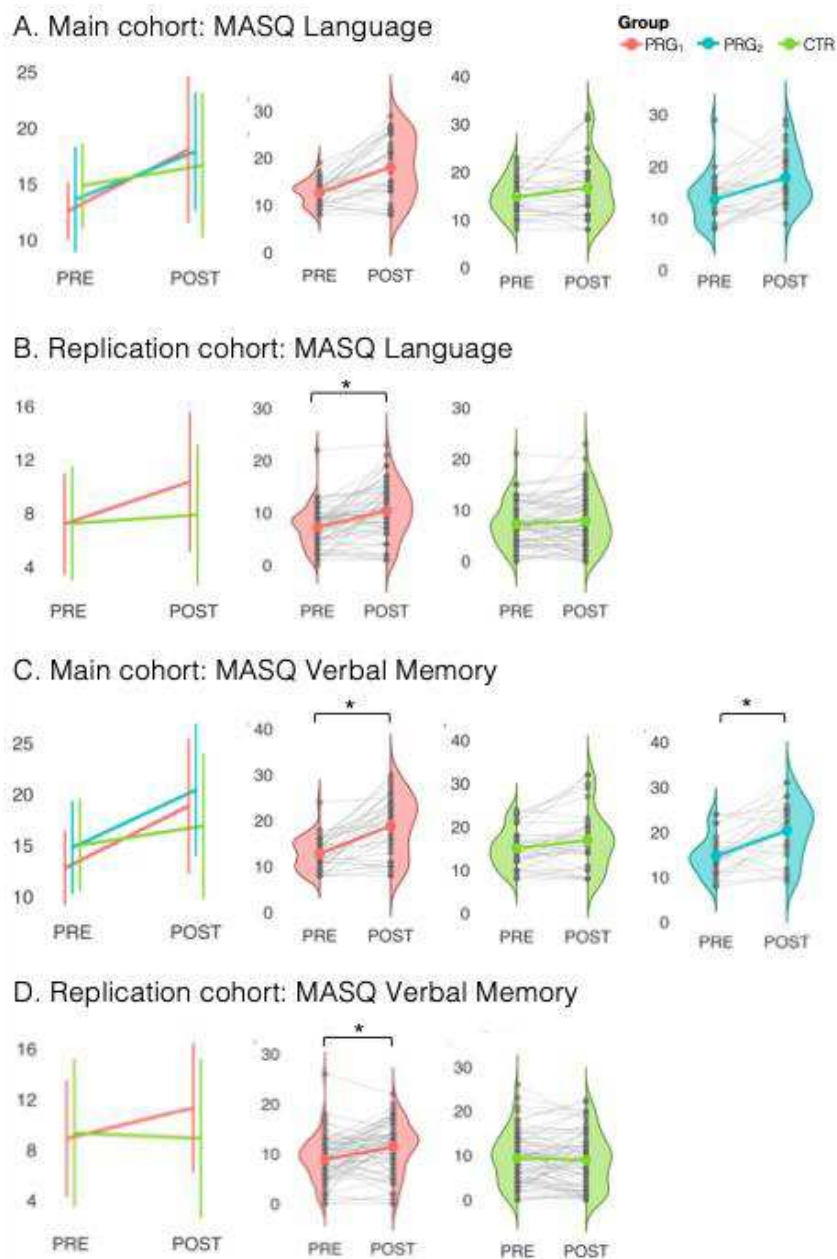
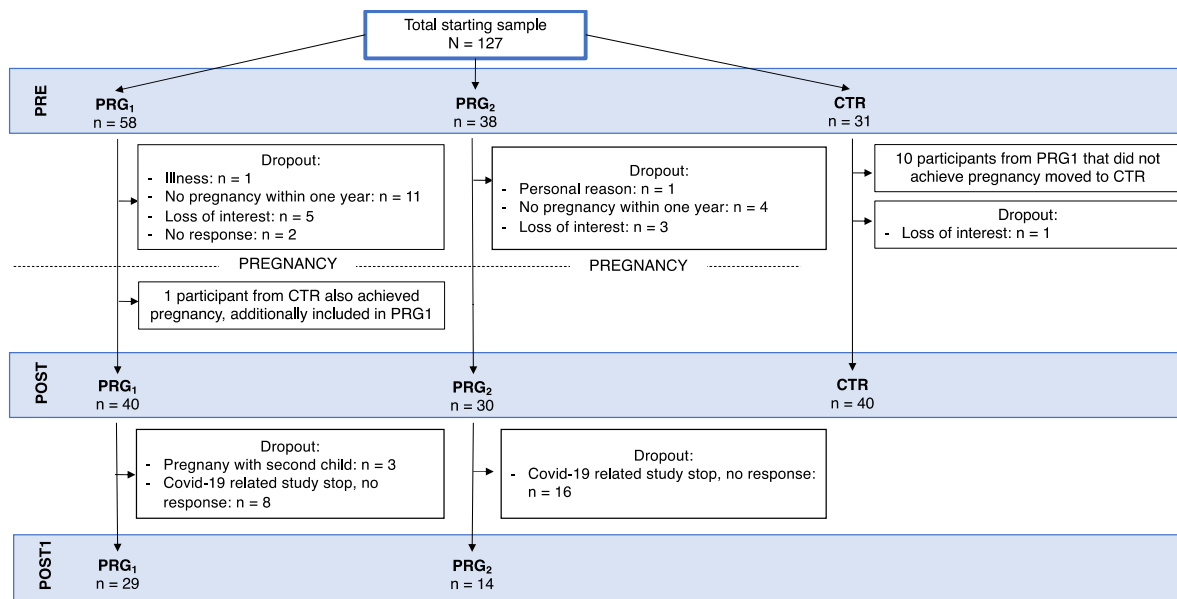


Table 1. Task effect of all groups at the baseline measure (PRE) during successful memory encoding. Voxel-wise FWE corrected $p < 0.05$ is presented.

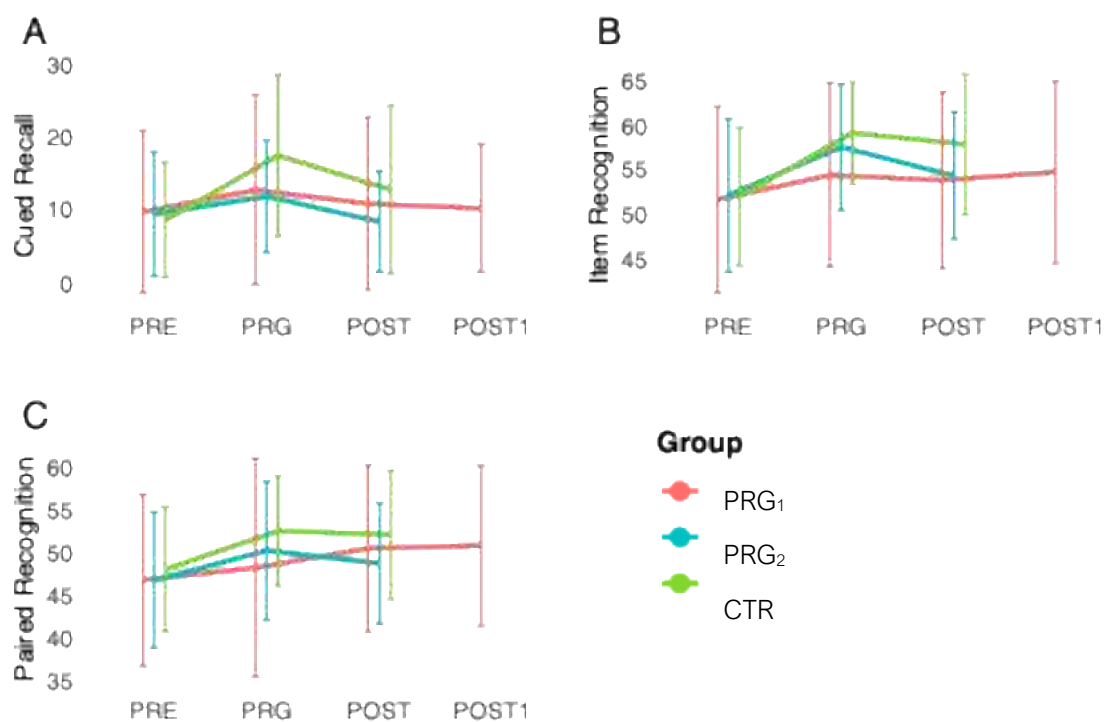
Regions	L/R	x	y	z	Voxel Size	F	p- value
Precuneus	R	12	-66	36	455	91.54	<0.001
Inferior frontal gyrus (orbital part)	L	-36	36	-12	460	76.78	<0.001
Supramarginal gyrus	R	57	-45	36	313	66.67	<0.001
Superior frontal gyrus	R	27	-51	21	269	52.41	<0.001
Posterior cingulate cortex	R	3	-39	21	137	48.17	<0.001
Middle occipital gyrus	L	-42	-72	24	73	41.28	<0.001
Superior frontal gyrus	L	-21	18	51	137	36.42	0.001
Middle temporal gyrus	L	-57	-45	-12	17	30.76	0.006
Inferior frontal gyrus (orbital gyrus)	R	30	36	-9	4	29.25	0.010
Caudate gyrus	L	-12	15	12	5	28.84	0.011
Parahippocampal	R	33	-3	-30	4	28.42	0.013
Caudate gyrus	R	12	18	6	8	28.35	0.013
Medial orbital frontal gyrus	L	-6	57	-6	1	24.94	0.042

1 Supplementary Figure 1. Overview of recruitment of all three groups for the main cohort.

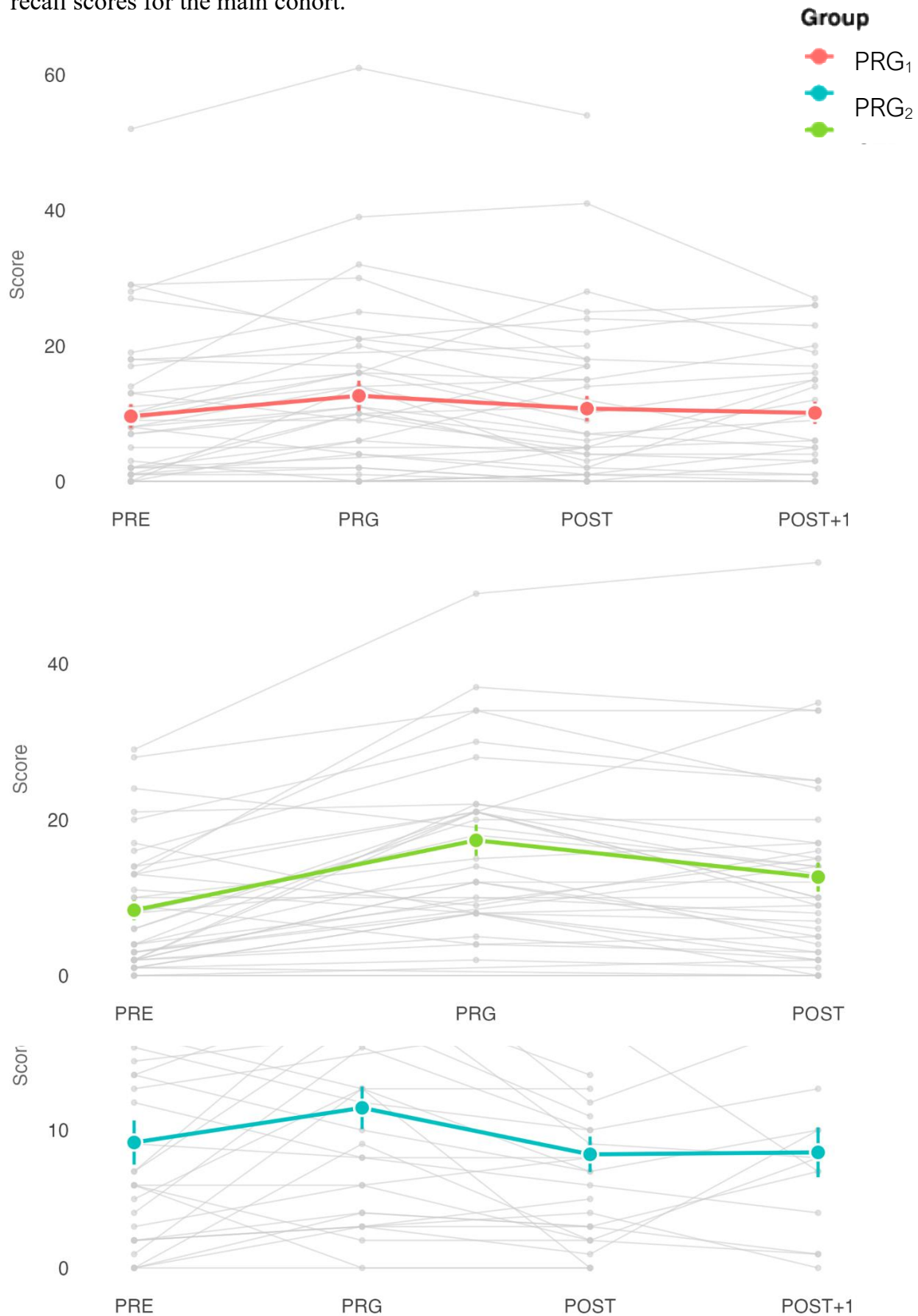


2

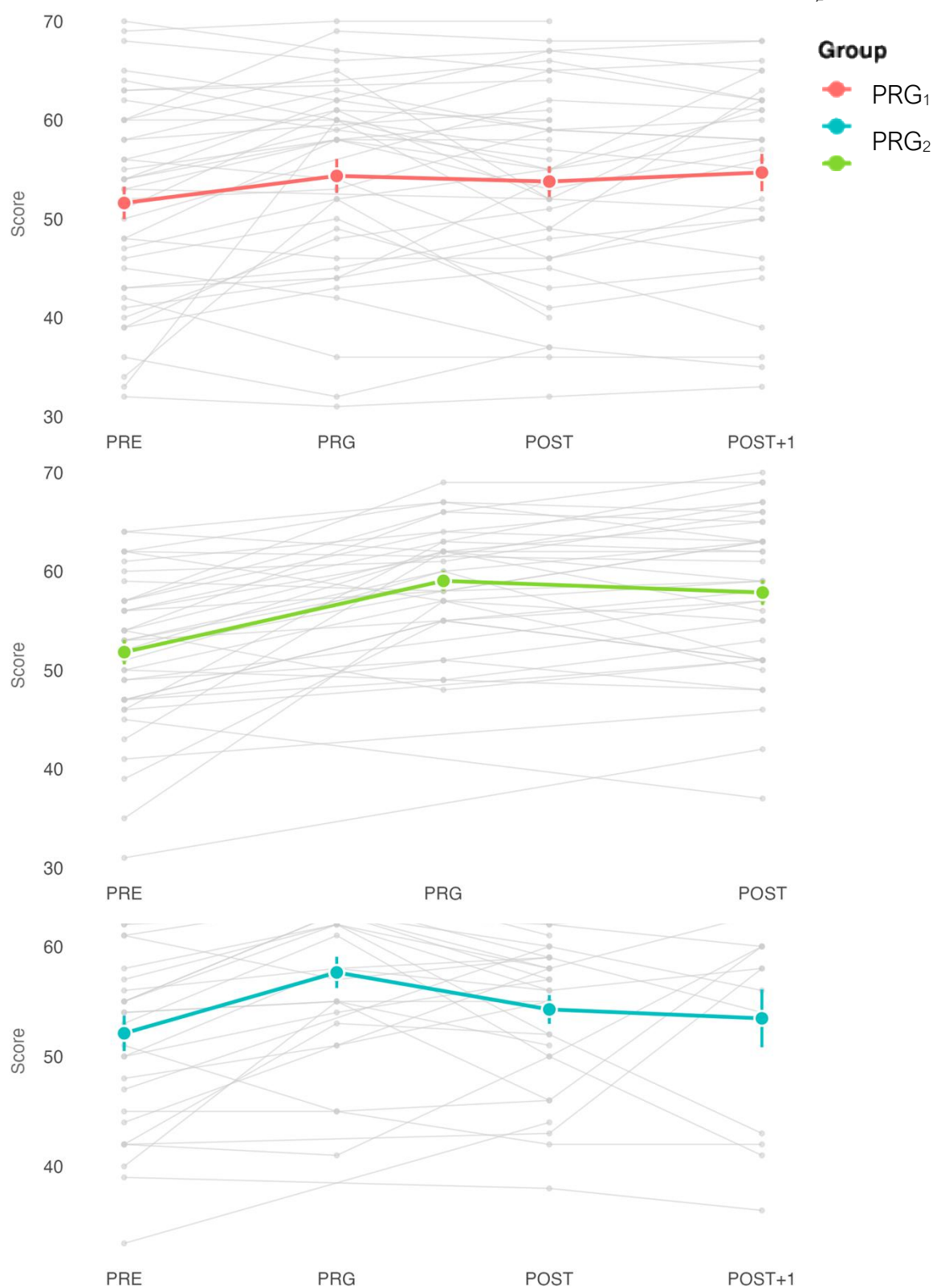
Supplementary Figure 2. Means and standard deviation for memory performance on the three tasks (A. Cued Recall, B. Item Recognition, C. Paired Recognition) for each group (first-time mothers, second-time mothers, nulliparous control women) and each timepoint (pre-conception, pregnancy, early postpartum and late postpartum) for the main cohort.



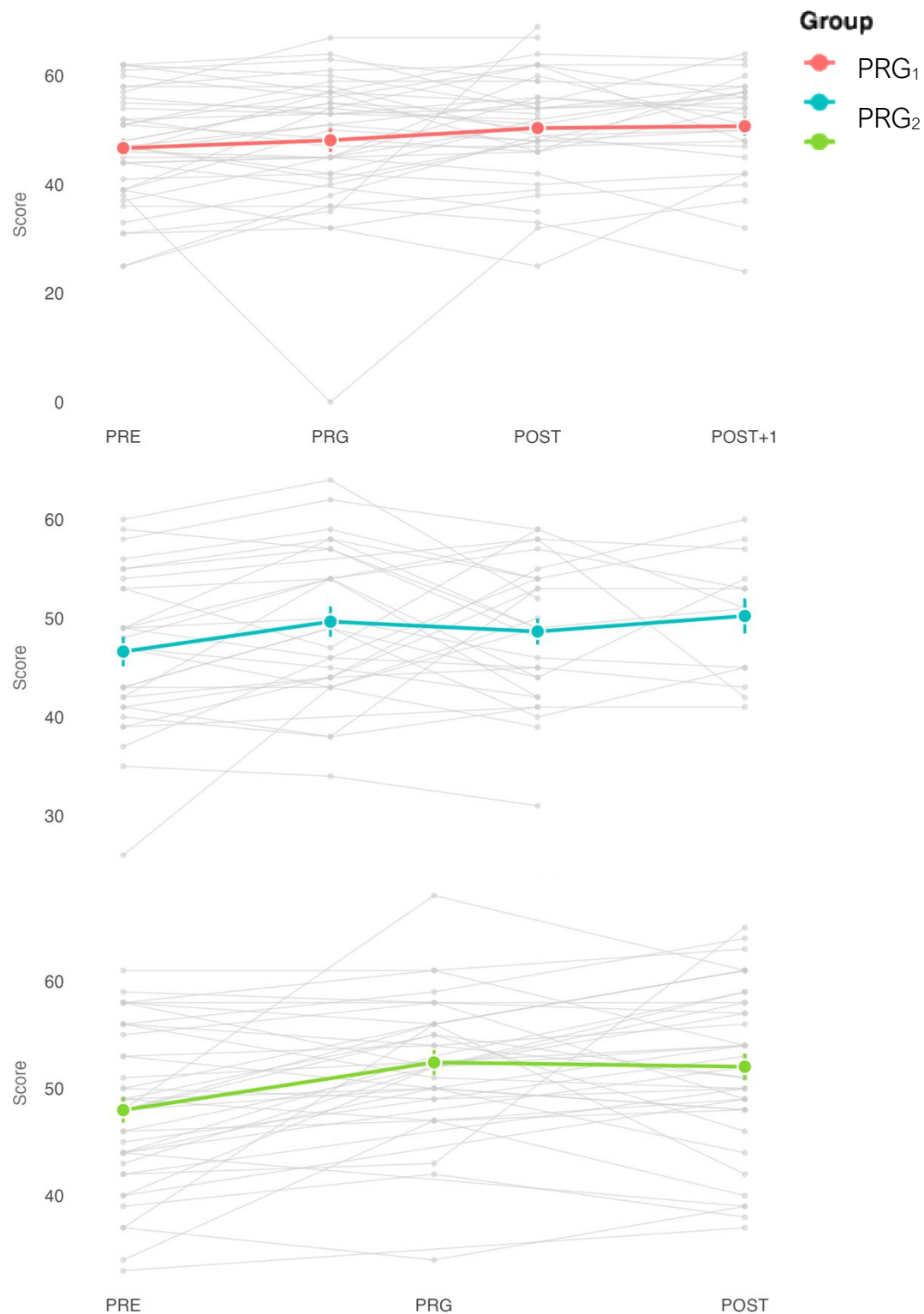
1 Supplementary Figure 3. Spaghetti plots for each group (first-time mothers; PRG₁, second-
 2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
 3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for cued
 4 recall scores for the main cohort.



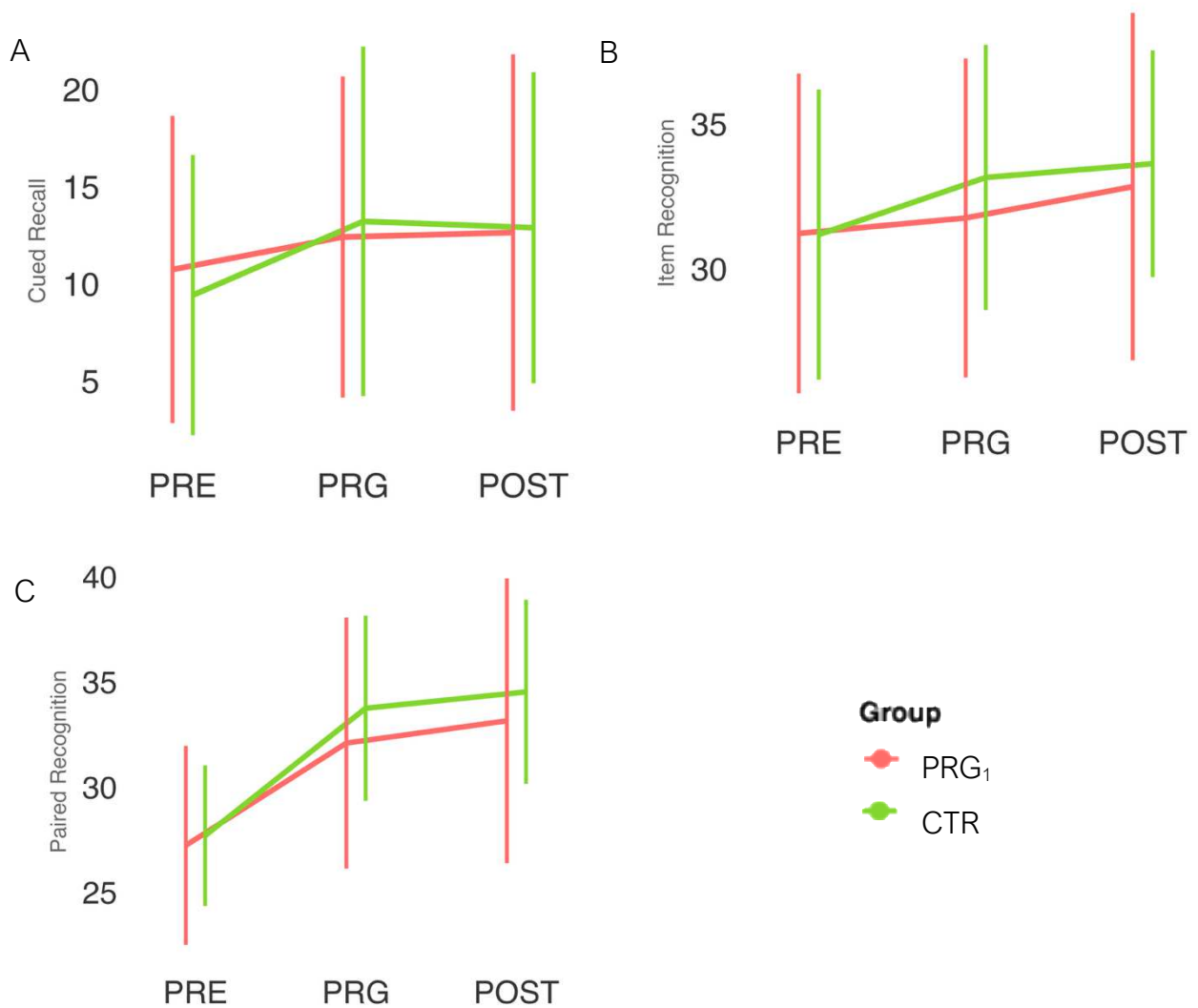
- 1 Supplementary Figure 4. Spaghetti plots for each group (first-time mothers; PRG₁, second-
- 2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
- 3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for item
- 4 recall scores for main cohort.



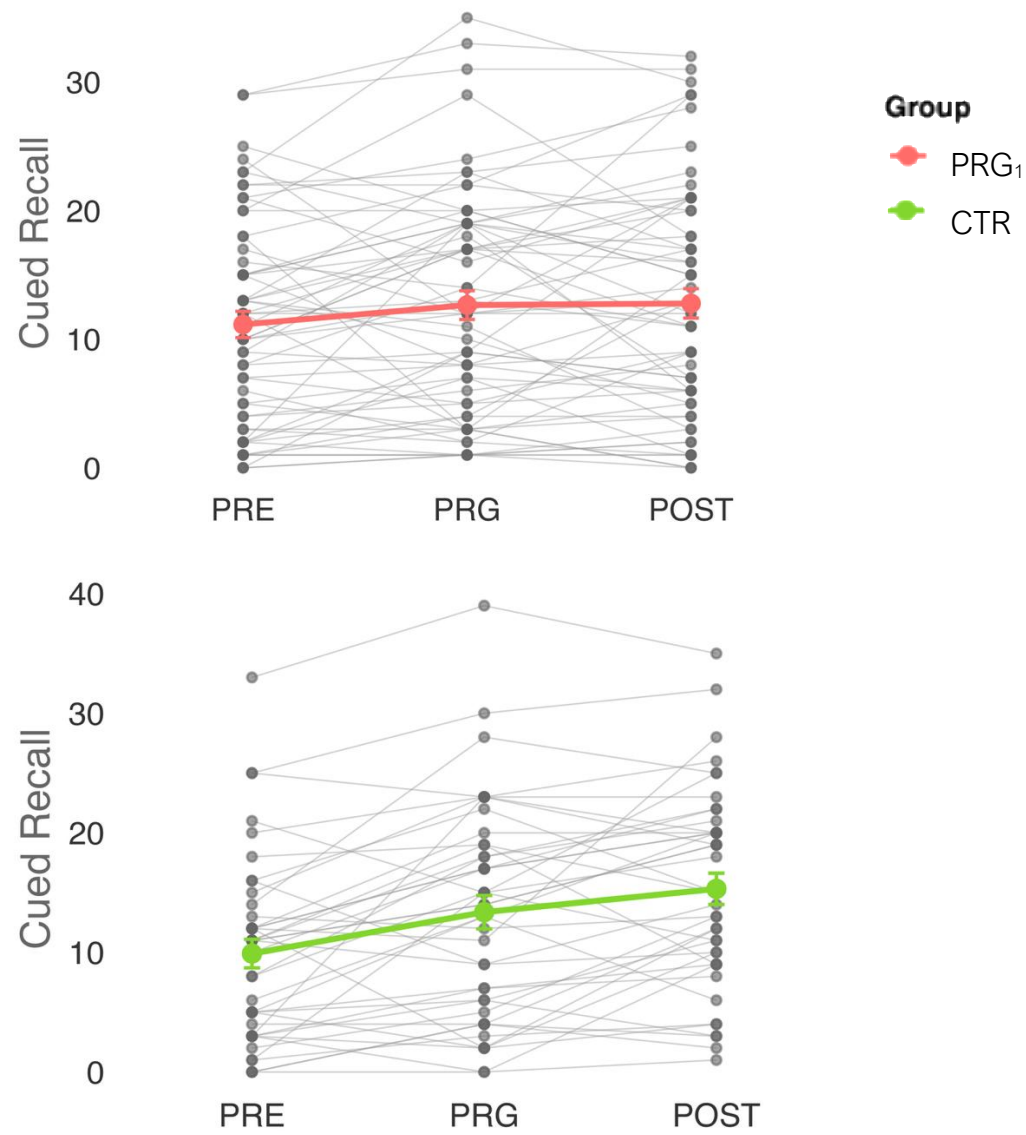
Supplementary Figure 5. Spaghetti plots for each group (first-time mothers; PRG₁, second-time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for paired recognition scores for the main cohort.



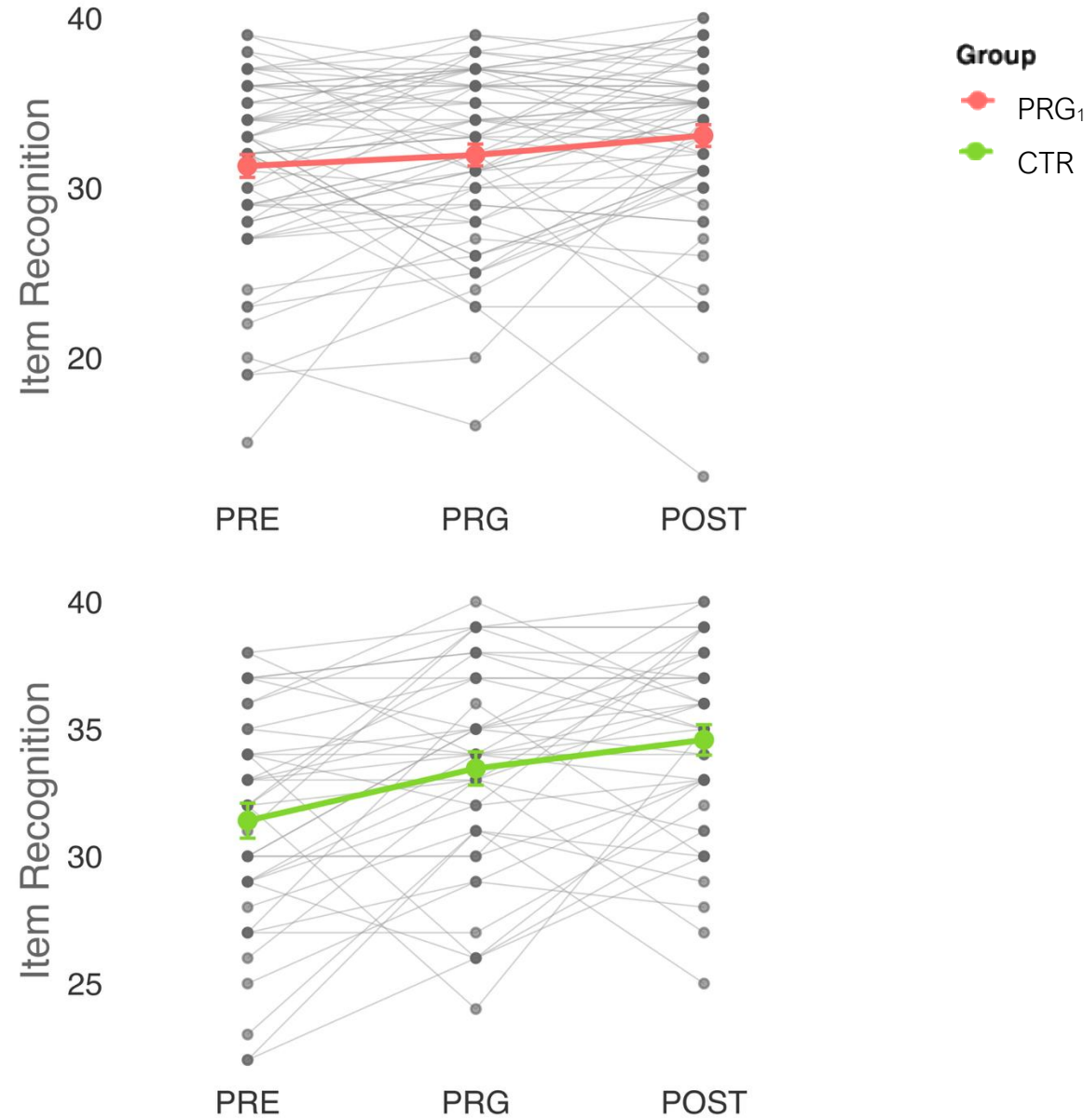
1 Supplementary Figure 6. Means and standard deviation for memory performance on the three
 2 tasks (A. Cued Recall, B. Item Recognition, C. Paired Recognition) for each group (first-time
 3 mothers, nulliparous control women) and each timepoint (pre-conception, pregnancy, early
 4 postpartum and late postpartum) for the replication cohort.



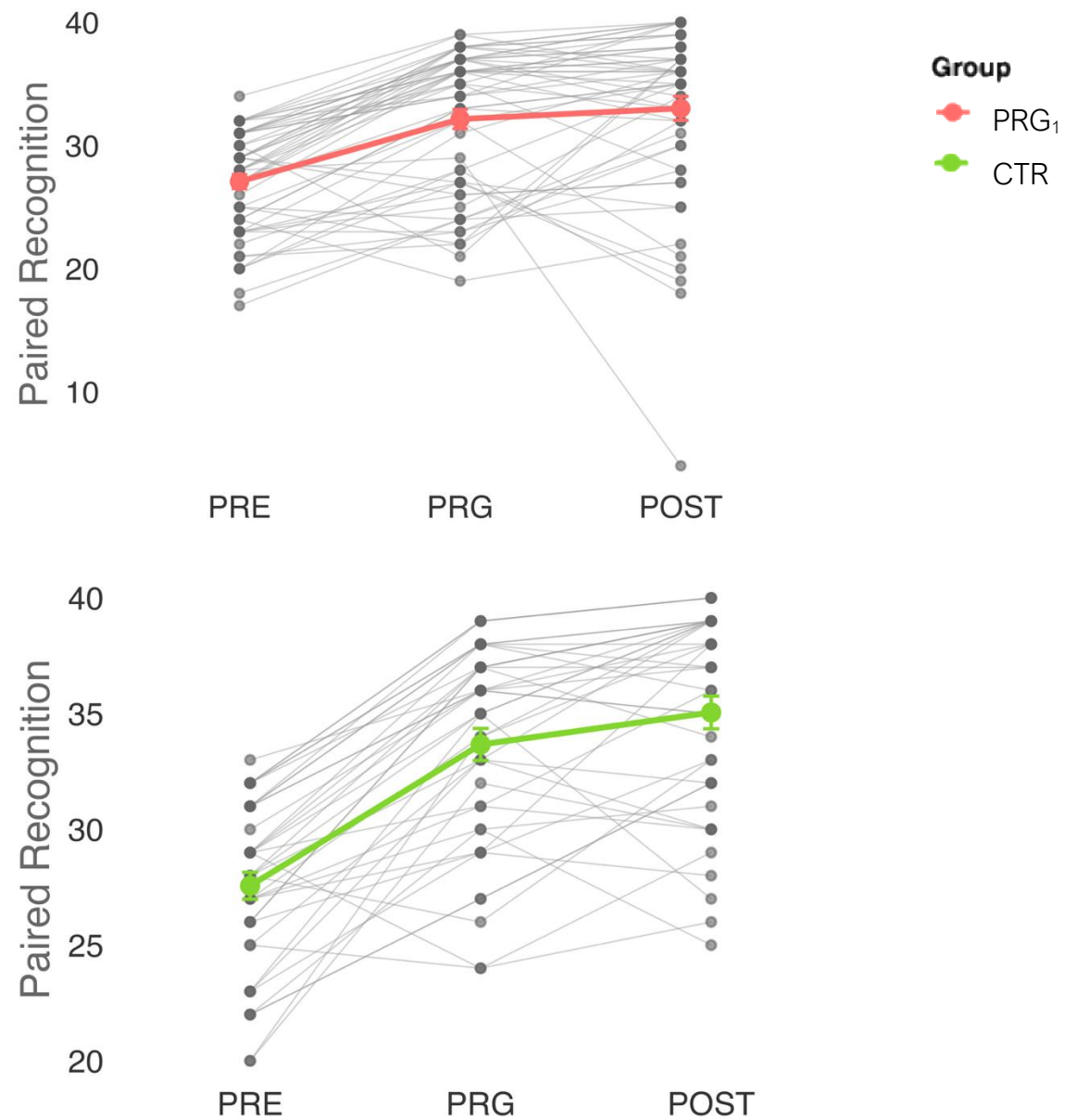
Supplementary Figure 7. Spaghetti plots for the replication cohort for each group (first-time mothers; PRG₁, nulliparous control women; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST) for cued recall for the replication cohort.



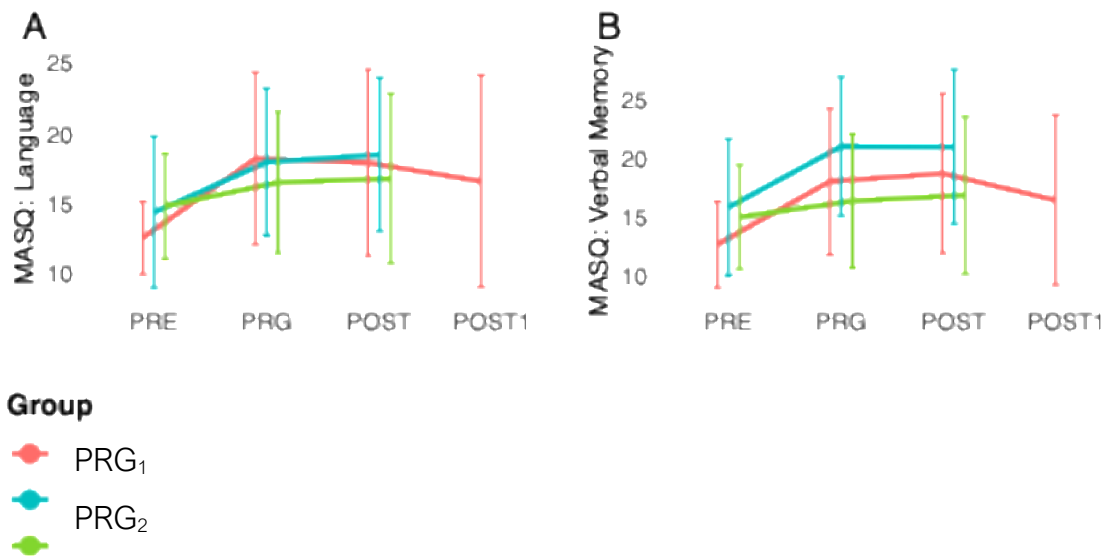
Supplementary Figure 8. Spaghetti plots for the replication cohort for each group (first-time mothers; PRG₁, nulliparous control women; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST) for item recognition for the replication cohort.



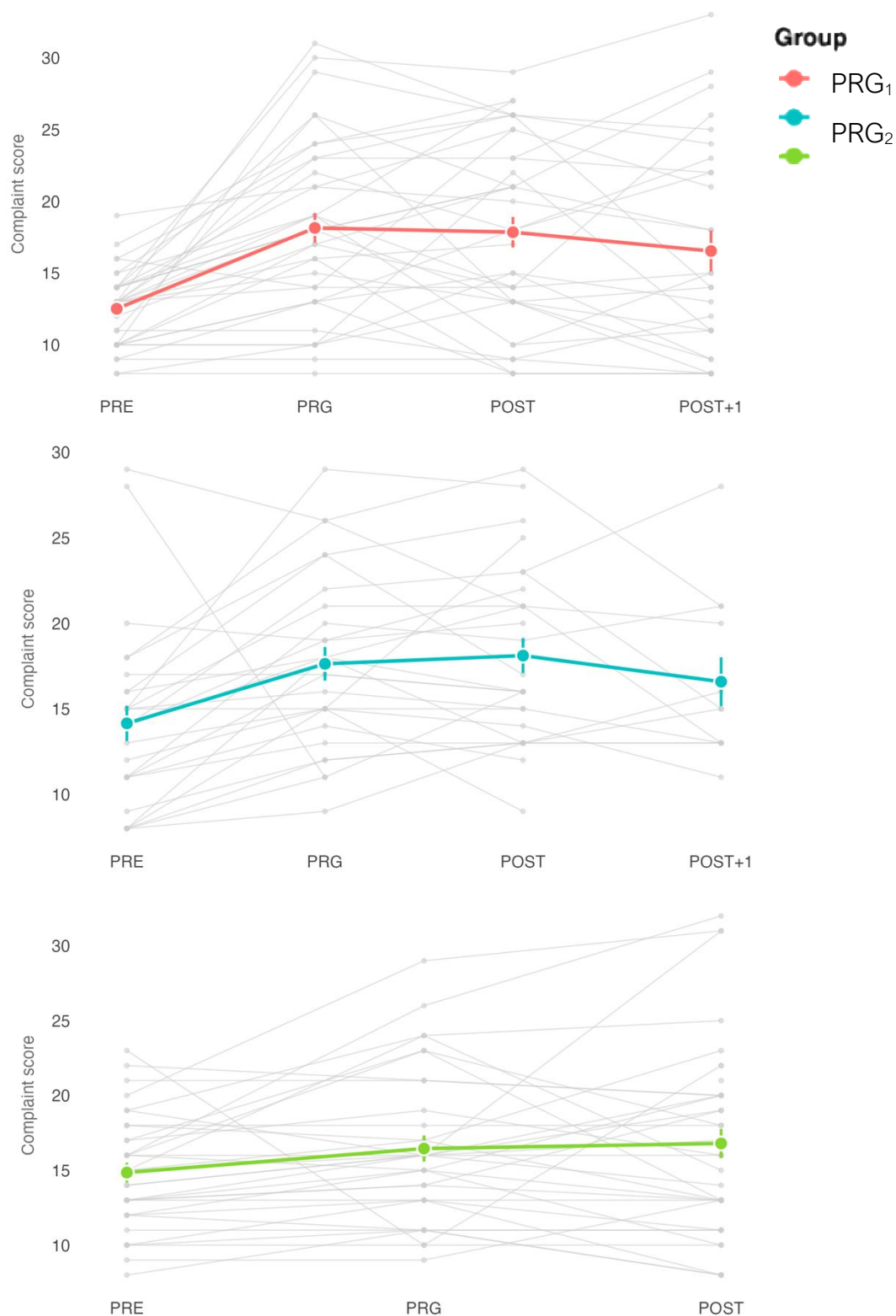
Supplementary Figure 9. Spaghetti plots for the replication cohort for each group (first-time mothers; PRG₁, nulliparous control women; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST) for paired recognition for the replication cohort.



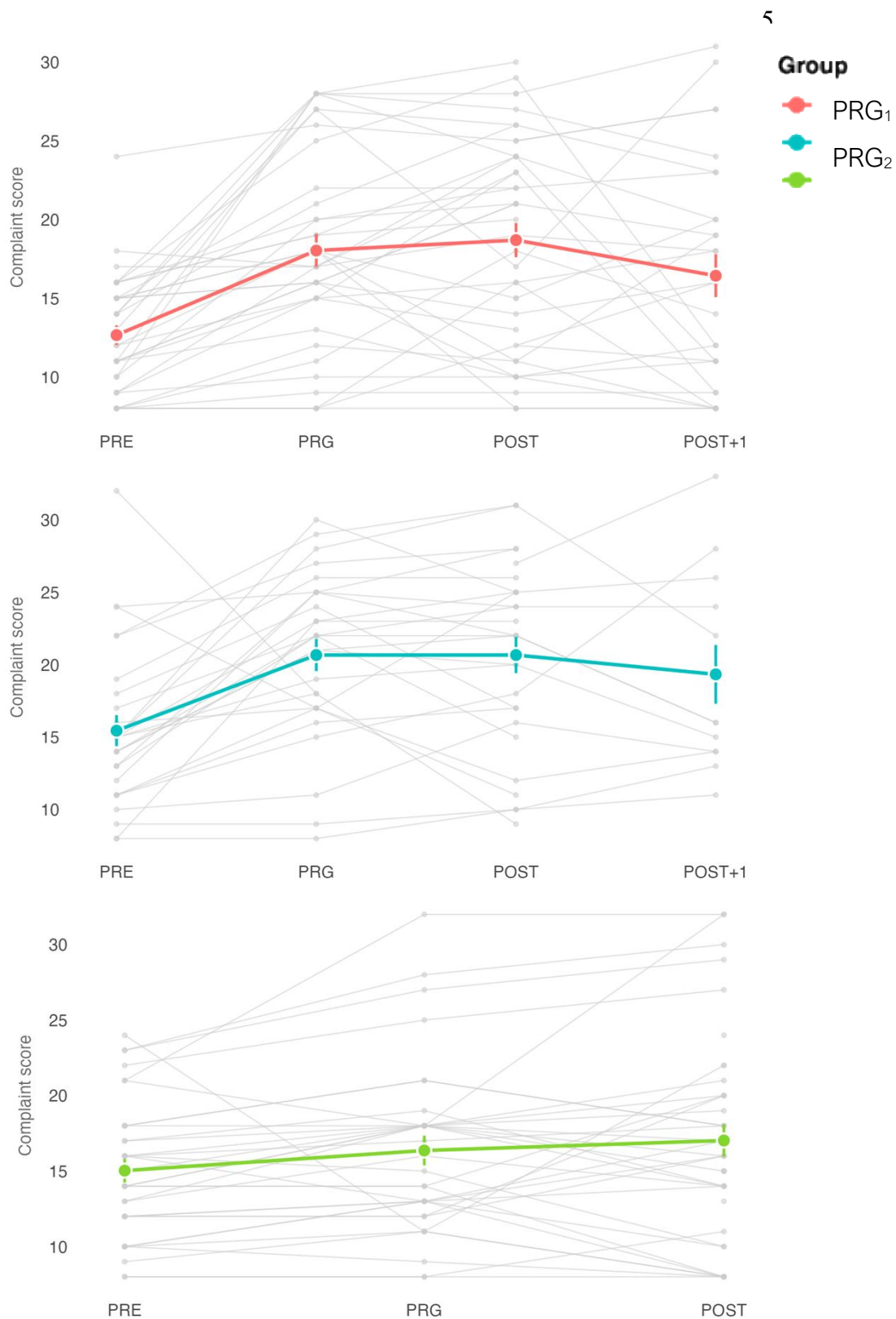
Supplementary Figure 10. Means and standard deviation for subjective memory deficits subscales A. Language and B. Verbal Memory for each group (first-time mothers; PRG₁, second-time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for the main cohort.



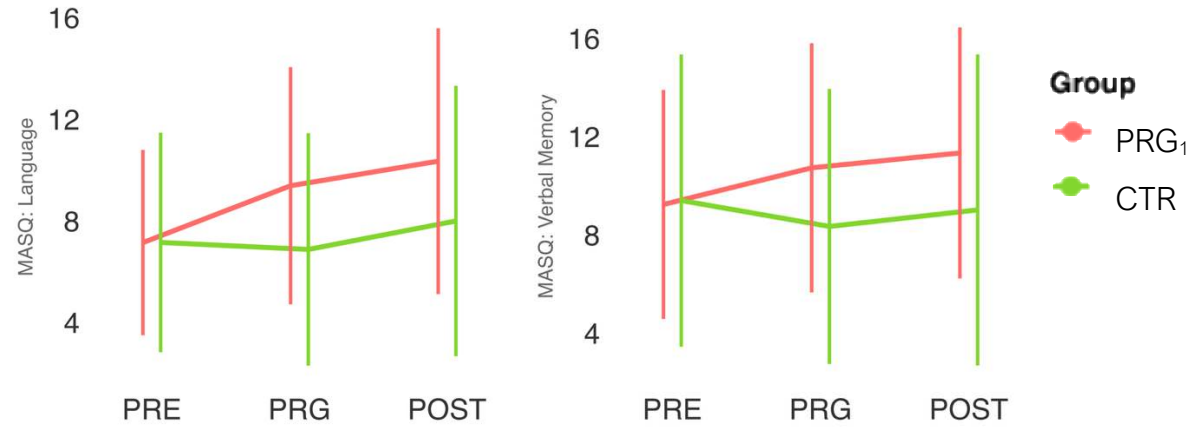
1 Supplementary Figure 11. Spaghetti plots for each group (first-time mothers; PRG₁, second-
2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for MASQ
4 Language subscale scores for the main cohort.



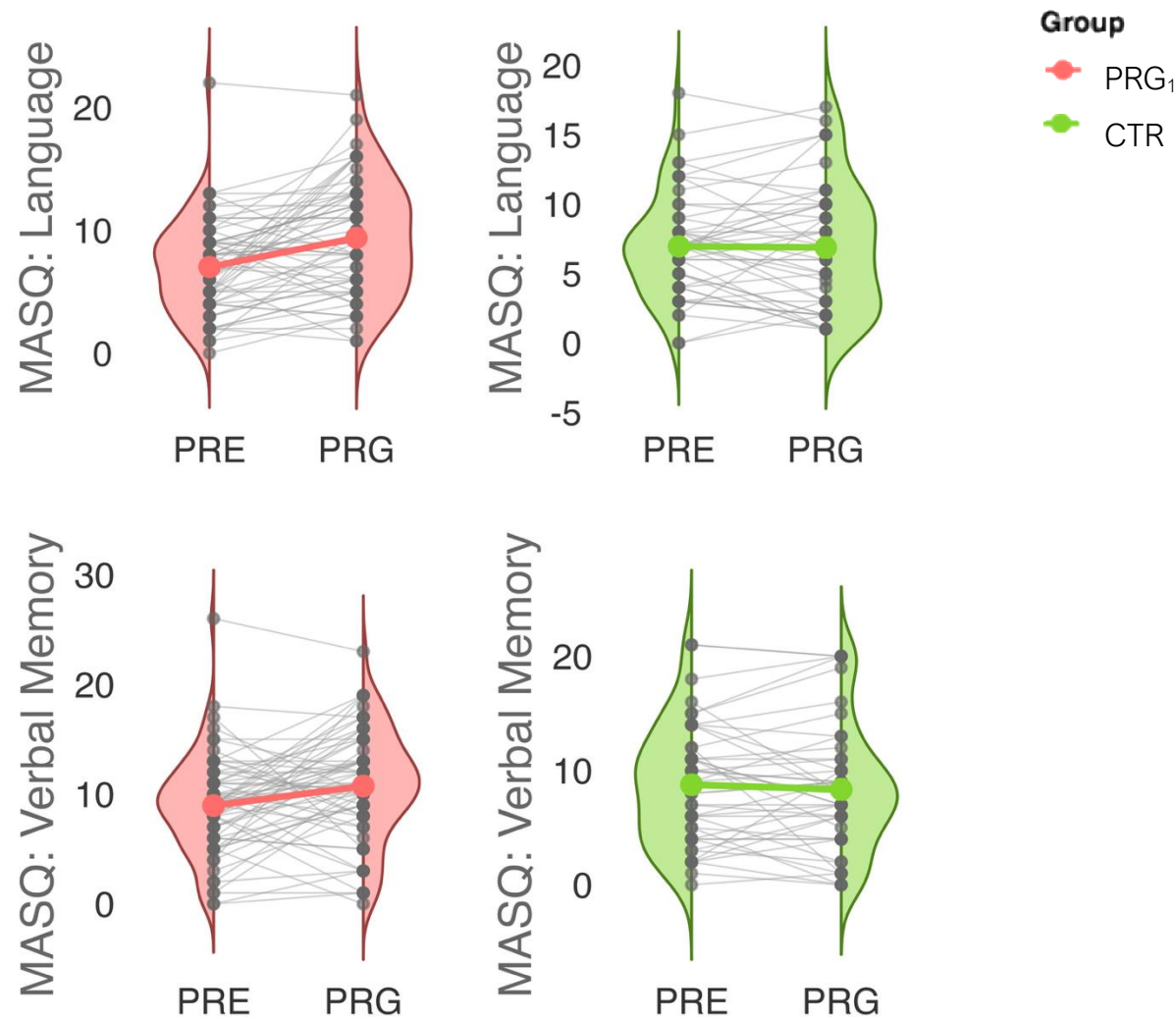
1 Supplementary Figure 12. Spaghetti plots for each group (first-time mothers; PRG₁, second-
2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for MASQ
4 Verbal Memory subscale scores for main cohort.



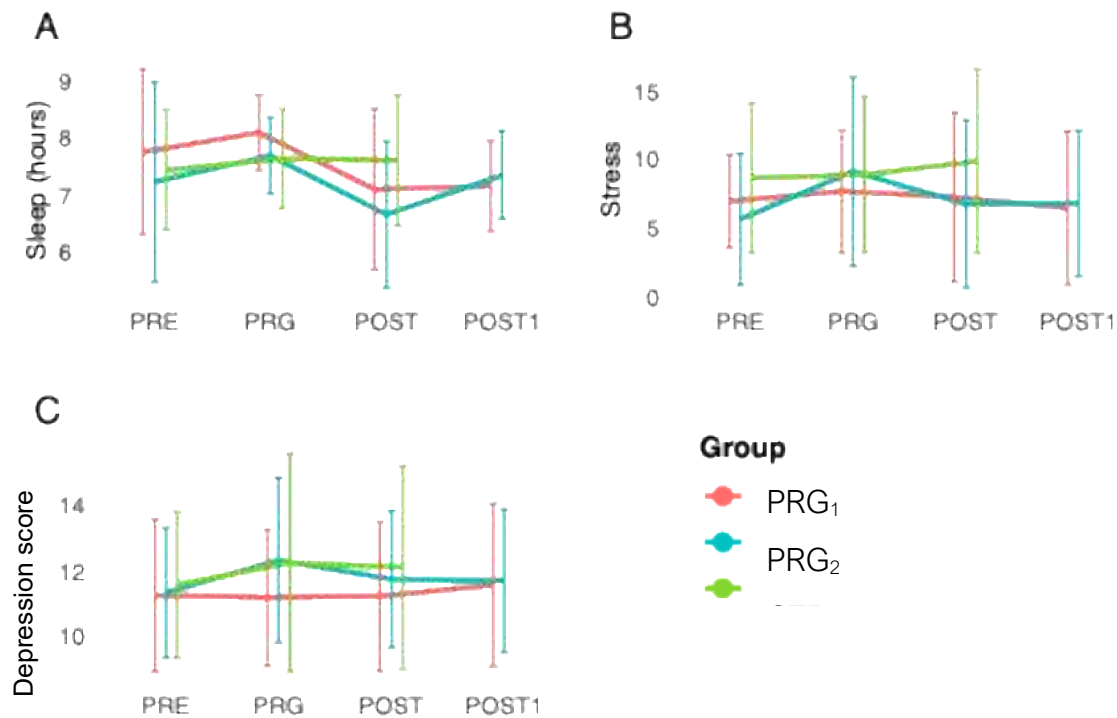
Supplementary Figure 13. Means and standard deviation for subjective memory deficits subscales A. Language and B. Verbal Memory for each group (first-time mothers; PRG₁, second-time mothers; CTR) and each timepoint (pre-conception; PRE, pregnancy; PRG, early postpartum; POST) for the replication cohort.



Supplementary Figure 14. Spaghetti plots for each group (first-time mothers; PRG₁, nulliparous control women; CTR) for the pre-conception (PRE) and pregnancy (PRG) timepoint for MASQ language and verbal memory subscale scores for the replication cohort.

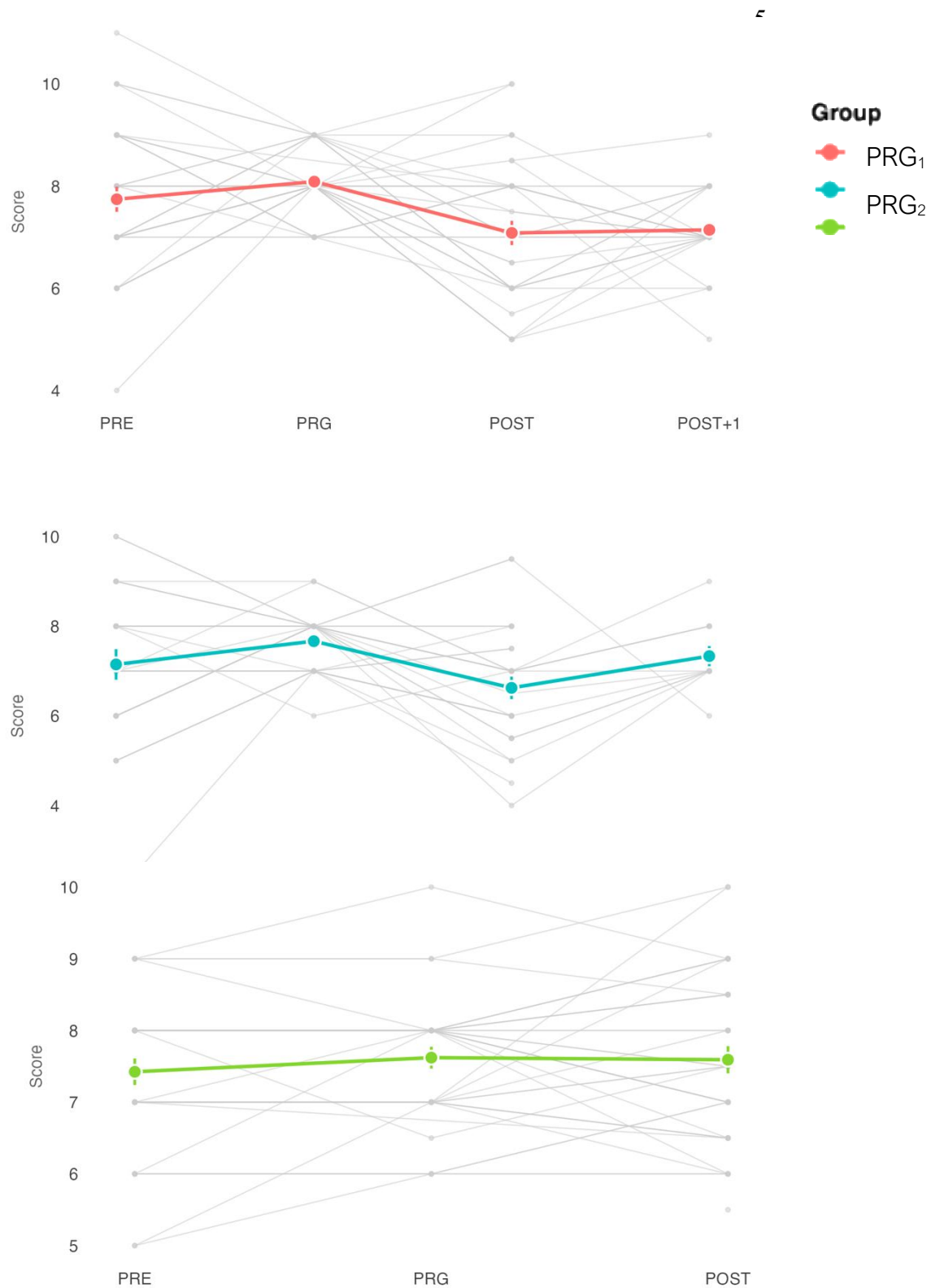


1 Supplement Figure 15. Means and standard deviation for A. hours of sleep, B. level of
 2 distress, and 3. depression scores for each group (first-time mothers, second-time mothers,
 3 nulliparous control women) and each timepoint (pre-conception, pregnancy, early postpartum
 4 and late postpartum) for the main cohort.

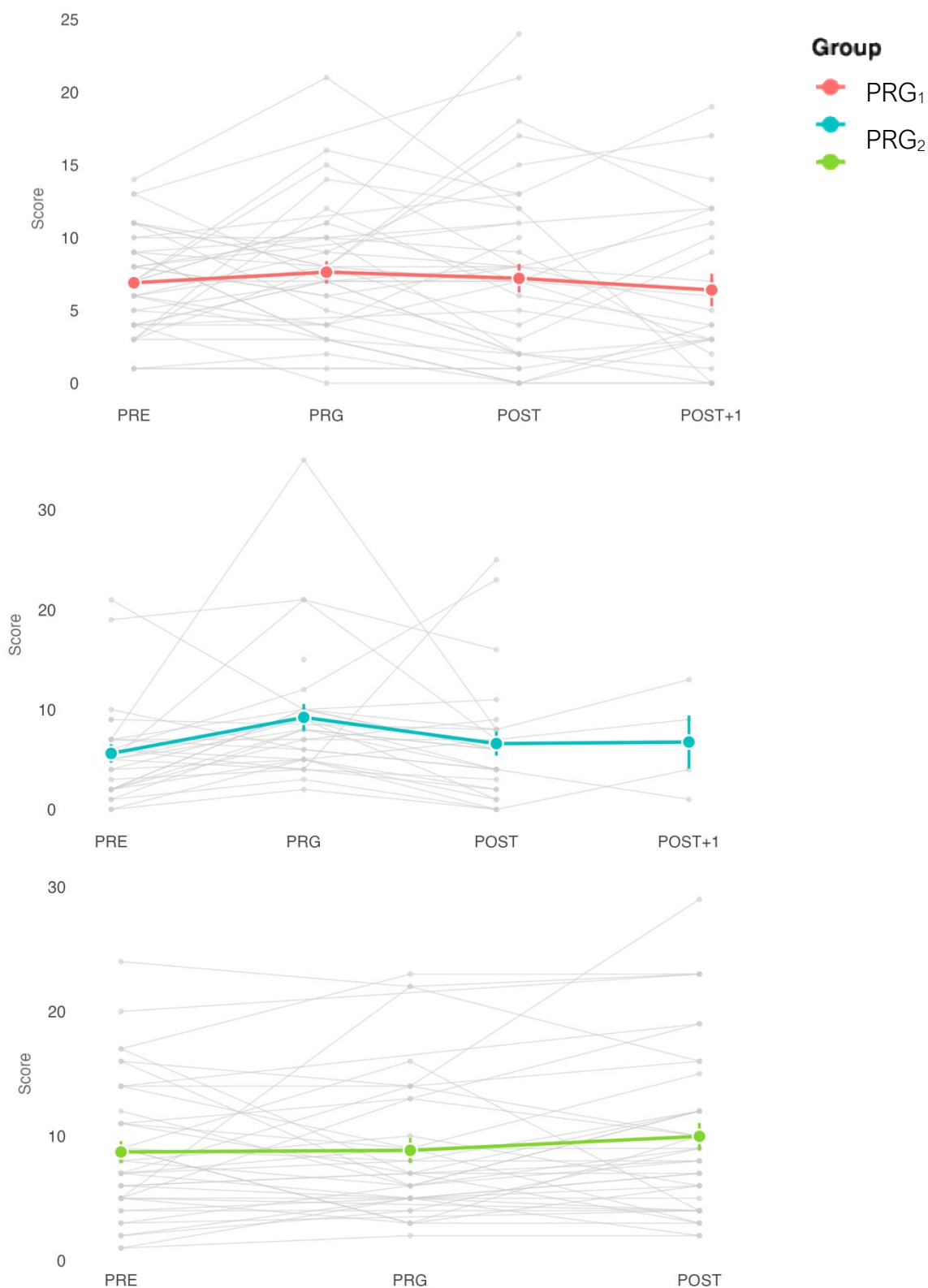


5

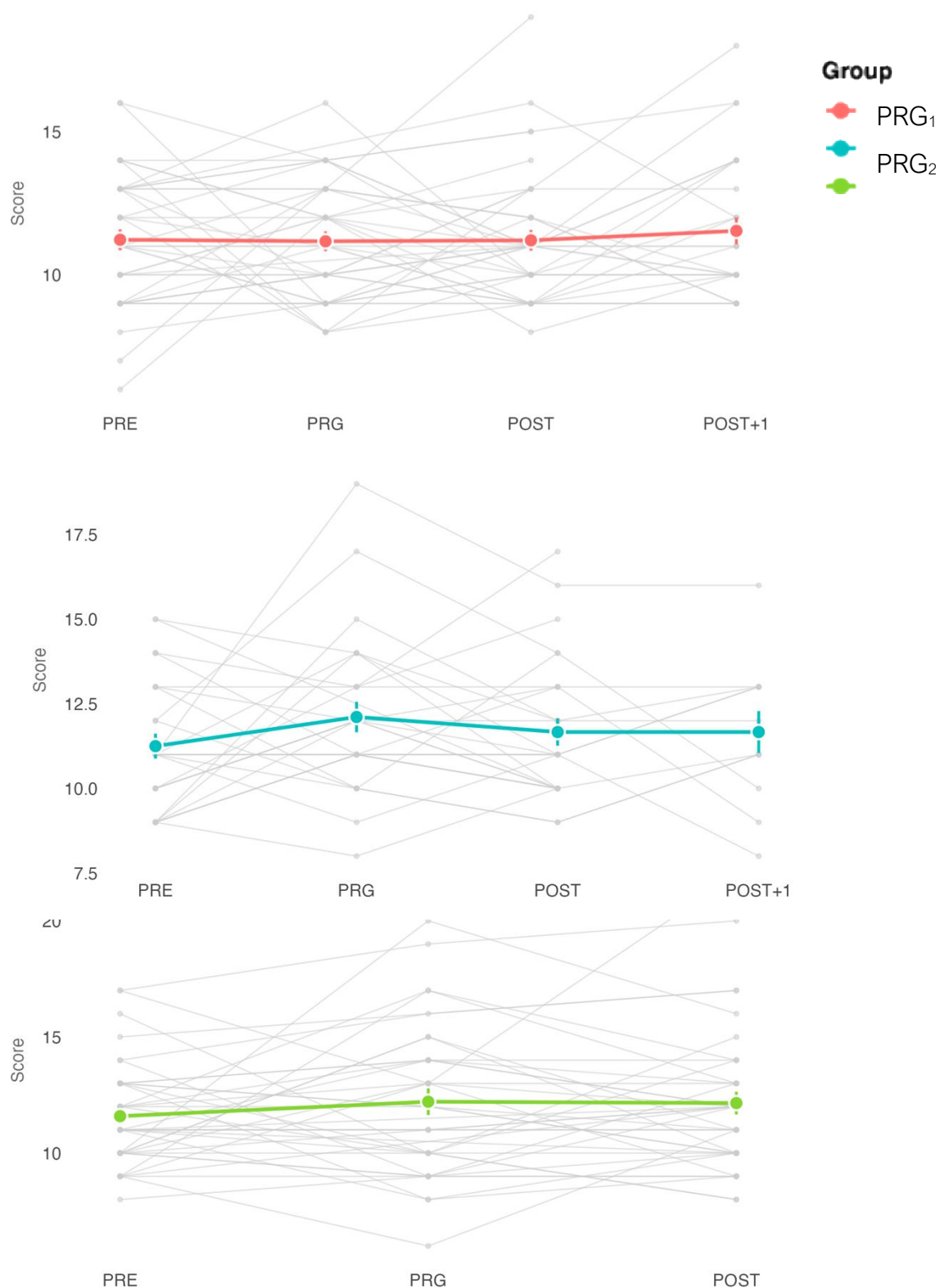
1 Supplementary Figure 16. Spaghetti plots for each group (first-time mothers; PRG₁, second-
 2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
 3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for Sleep
 4 (hours average past week).



1 Supplementary Figure 17. Spaghetti plots for each group (first-time mothers; PRG₁, second-
 2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
 3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for
 4 psychological distress measured with K10 questionnaire.



1 Supplementary Figure 18. Spaghetti plots for each group (first-time mothers; PRG₁, second-
 2 time mothers; PRG₂, nulliparous control women; CTR) and each timepoint (pre-conception;
 3 PRE, pregnancy; PRG, early postpartum; POST and late postpartum; POST+1) for
 4 depression measured with EPDS questionnaire.



1 Supplementary Table 1. Demographic information for all groups for each group (first-time
2 mothers; PRG₁, second-time mothers; PRG₂, nulliparous control women; CTR).

Characteristic	PRG ₁	PRG ₂	CTR
Age at PRE (M ± SD)	29.35 ± 3.51	32.03 ± 2.33	29.33 ± 3.57
Education (M Verhage score ± SD)	6.43 ± 0.78	6.33 ± 0.76	6.65 ± 0.53
Secondary school	-	2	-
College	19	13	12
University	21	14	28
Unknown	-	1	-
Days PRE - POST (M ± SD)	509.78 ± 158.91	502.7 ± 125.02	457.95 ± 81.95
Days parturition - POST (M ± SD)	100.6 ± 70.8	95.4 ± 70.2	-
Days parturition - POST+1 (M ± SD)	402.6 ± 64.8	-	-
Medical History			
ADHD	1	-	-
Burn-out	-	-	1
Depression/anxiety	3	3	5
Eating disorder	1	2	-
Harlequin syndrome	1	-	-
Meningitis	1	-	-
Neuralgic amyotrophy	-	1	-
Personality disorder	-	1	-
Trigeminal neuralgia	-	-	1
Natural Conception	37	29	-
Assisted Conception	3	1	-
Vaginal Birth	31	28	-
Caesarean section	9	2	-
Twins	1	-	-
Daughters	23	15	-
Sons	17	13	-
Sex baby unknown	1	2	-
Feeding method at POST session			-
Breastfeeding	30	21	-
Formula	10	7	-
NA	-	2	-

Supplementary Table 2. Brain-behavior correlations between memory encoding-related neural activity and post-minus-pre change scores in memory performance in first-time mothers (PRG₁), second-time mothers (PRG₂), and nulliparous controls (CTR). Beta-weights were extracted from the associative memory encoding network as defined by Kim (2024) using a region-of-interest mask. Pearson or Spearman correlations were selected based on normality of the data (Shapiro-Wilk test).

Group	Outcome	n	Method	r/rho	T	CI	p	p _{FDR}
PRG ₁	cued recall	30	spearman	0.39	2743.69		0.033	0.05
	item recognition	30	pearson	0.46	2.77	[0.1247 0.7063]	0.0098	0.029
	paired recognition	30	spearman	0.049	4284.65		0.81	0.81
PRG ₂	cued recall	26	pearson	-0.38	-2.00	[-0.668 0.0102]	0.057	0.17
	item recognition	26	pearson	-0.091	-0.45	[-0.4625 0.3066]	0.66	0.66
	paired recognition	26	pearson	-0.20	-0.99	[-0.5433 0.2056]	0.33	0.50
CTR	cued recall	32	spearman	0.0020	5444.96		0.99	0.99
	item recognition	32	spearman	0.049	5186.21		0.79	0.99
	paired recognition	32	spearman	-0.040	5676.38		0.83	0.99

Note: PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy; PRG = during pregnancy, POST = early postpartum, POST+1 = late postpartum. P_{FDR} = p-value corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate procedure, applied separately per group across the three memory outcomes. CI = 95% confidence interval (Pearson only).

Supplementary Table 3. Brain-behavior correlations between memory encoding-related neural activity and post-minus-pre change scores in subjective scores (MASQ Verbal Memory and Language subscales) in first-time mothers (PRG₁), second-time mothers (PRG₂), and nulliparous controls (CTR). Beta-weights were extracted from the associative memory encoding network as defined by Kim (2024) using a region-of-interest mask. Pearson or Spearman correlations were selected based on normality of the data (Shapiro-Wilk test).

Group	Outcome	n	Method	r/rho	T	CI	p	p _{FDR}
PRG ₁	Language	30	pearson	0.18	0.97	[-0.19 0.51]	0.34	0.68
	Verbal	30	pearson	0.071	0.37	[-0.30 0.42]	0.71	0.71
	Memory							
PRG ₂	Language	23	pearson	0.024	0.11	[-0.39 0.43]	0.91	0.91
	Verbal	23	pearson	0.037	0.17	[-0.38 0.44]	0.87	0.91
	Memory							
CTR	Language	27	spearman	0.14	2807.78		0.48	0.85
	Verbal	27	spearman	0.039	3148.38		0.85	0.85
	Memory							

Note: PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. P_{FDR} = p-value corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate procedure, applied separately per group across the three memory outcomes. CI = 95% confidence interval (Pearson only).

Supplementary Table 4. Means and standard deviations for the memory performance score per task and the two subscales of the MASQ per group for each group and each timepoint for the main cohort.

Behavioral measure	Group	PRE	PRG	POST	POST+1
Cued Recall	PRG ₁	9.6 ± 11.18	12.63 ±	10.72 ±	10.10 ± 8.83
			13.14	11.84	
	CTR	8.45 ± 7.94	17.35 ±	12.65 ±	-
			11.17	11.58	
	PRG ₂	9.3 ± 8.53	11.7 ± 7.73	8.24 ± 6.95	-
Item Recognition	PRG ₁	51.6 ± 10.36	54.34 ±	53.77 ± 9.81	54.69 ±
			10.22		10.13
	CTR	51.95 ± 7.67	59.03 ± 5.68	57.77 ± 7.76	-
	PRG ₂	52.1 ± 8.55	57.46 ± 7.02	54.28 ± 7.05	-
Paired Recognition	PRG ₁	46.7 ± 10.02	48.14 ±	50.4 ± 9.73	50.72 ± 9.35
			12.73		
	CTR	47.95 ± 7.25	52.42 ± 6.43	52 ± 7.52	-
	PRG ₂	46.67 ± 7.9	50.15 ± 8.08	48.66 ± 7.07	-
MASQ: Language	PRG ₁	12.51 ± 2.61	18.14 ± 6.11	17.85 ± 6.63	16.54 ± 7.51
	CTR	14.76 ± 3.77	16.45 ± 5.04	16.73 ± 6.02	-
	PRG ₂	14.15 ± 5.39	17.63 ± 5.12	18.11 ± 5.28	-
MASQ: Verbal Memory	PRG ₁	12.66 ± 3.69	18.03 ± 6.21	18.69 ± 6.79	16.43 ± 7.22
	CTR	15.00 ± 4.44	16.36 ± 5.67	16.85 ± 6.69	-
	PRG ₂	15.82 ± 5.8	21.00 ± 5.93	20.96 ± 6.57	-

Note: PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy; PRG = during pregnancy, POST = early postpartum, POST+1 = late postpartum

Supplementary Table 5. Repeated Measures ANCOVA results for PRE, and POST
timepoints across all three groups for Cued Recall memory performance for the main cohort.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	187.45	0.09	0.91	0.0020
Session	1	81	25.83	5.38	0.023	0.062
Group $\times$ Session	2	81	25.83	2.58	0.082	0.060
Sleep Δ	1	81	187.45	0.10	0.75	0.0010
K10 Δ	1	81	187.45	2.37	0.13	0.028
Edinburgh Δ	1	81	187.45	2.82	0.097	0.034
Age	1	81	187.45	1.26	0.27	0.015

Note: Δ = change score (POST-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 6. Repeated Measures ANCOVA results for PRE, and POST timepoints across all three groups for Item Recognition memory performance for the main cohort.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	105.13	1.54	0.22	0.037
Session	1	81	22.84	19.26	< 0.001	0.19
Group $\times$ Session	2	81	22.84	4.02	0.022	0.090
Sleep Δ	1	81	105.13	0.69	0.41	0.0090
K10 Δ	1	81	105.13	5.19	0.025	0.060
Edinburgh Δ	1	81	105.13	3.15	0.080	0.037
Age	1	81	105.13	0.08	0.78	0.001

Note: Δ = change score (POST-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 7. Repeated Measures ANOVA results for PRE and POST timepoints across all three groups for Paired Recognition memory performance for the main cohort.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	114.84	1.73	0.18	0.041
Session	1	81	40.57	16.17	< 0.001	0.17
Group $\times$ Session	2	81	40.57	0.30	0.74	0.0070
Sleep Δ	1	81	114.84	0.44	0.51	0.0050
K10 Δ	1	81	114.84	0.82	0.37	0.010
Edinburgh Δ	1	81	114.84	0.12	0.73	0.0020
Age	1	81	114.84	0.47	0.49	0.0060

Note: Δ = change score (POST-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 8. Pearson correlations between POST-PRE memory difference scores across the full sample for the main cohort.

	Cued Recall	Item Recognition	Paired Recognition
Cued Recall	-	$r = 0.60, p < 0.001$	$r = 0.37, p < 0.001$
Item Recognition	$r = 0.60, p < 0.001$	-	$r = 0.37, p < 0.001$
Paired Recognition	$r = 0.37, p < 0.001$	$r = 0.37, p < 0.001$	-

Note. Correlations above and below the diagonal are identical.

Supplementary Table 9. Means and standard deviations for the memory performance score per task and the two subscales of the MASQ per group for each group and each timepoint for replication cohort.

Behavioral measure	Group	PRE	PRG	POST
Cued Recall	PRG ₁	10.74 ± 7.90	12.41 ± 8.26	12.64 ± 9.17
	CTR	9.42 ± 7.21	13.21 ± 8.99	12.89 ± 8.00
Item Recognition	PRG ₁	31.22 ± 5.52	31.75 ± 5.50	32.83 ± 6.00
	CTR	31.18 ± 5.01	33.15 ± 4.58	33.62 ± 3.91
Paired Recognition	PRG ₁	27.23 ± 4.73	32.09 ± 5.97	33.15 ± 6.77
	CTR	27.69 ± 3.34	33.75 ± 4.40	34.53 ± 4.38
MASQ: Language	PRG ₁	7.15 ± 3.65	9.28 ± 4.39	10.49 ± 5.02
	CTR	7.16 ± 4.32	6.84 ± 4.63	7.69 ± 5.39
MASQ: Verbal Memory	PRG ₁	9.24 ± 4.66	10.85 ± 4.81	11.22 ± 5.13
	CTR	9.40 ± 5.95	8.15 ± 5.73	8.76 ± 6.44

Supplementary Table 10. Repeated Measures ANCOVA results for PRE, and PRG timepoints across all three groups for Cued Recall memory performance.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	209.78	0.64	0.528	0.016
Session	1	81	21.05	43.54	< 0.001	0.350
Group $\times$ Session	2	81	21.05	9.25	< 0.001	0.186
Sleep Δ	1	81	209.78	0.27	0.607	0.003
K10 Δ	1	81	209.78	0.62	0.433	0.008
Edinburgh Δ	1	81	209.78	0.64	0.427	0.008
Age	1	81	209.78	0.31	0.580	0.004

Note: Δ = change score (PRG-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 11. Post-hoc for repeated Measures ANCOVA results for PRE and PRG timepoints across all three groups for cued recall memory performance scores.

Group	Session	Est.	SE	df	p_{all}	t
PRG ₁	PRE - PRG	-0.97	1.14	81	0.0028	-3.09
PRG ₂	PRE - PRG	-1.33	1.40	81	0.34	-0.95
CTR	PRE - PRG	-8.90	1.22	81	< 0.001	-7.29

Note: p_{all} = Tukey HSD adjusted p-value. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy, POST = early postpartum.

Supplementary Table 12. Repeated Measures ANCOVA results for PRE, and PRG timepoints across all three groups for Item Recognition memory performance.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	79	129.83	1.86	0.16	0.045
Session	1	79	20.41	43.79	< 0.001	0.36
Group $\times$ Session	2	79	20.41	2.17	0.12	0.052
Sleep Δ	1	79	129.83	0.05	0.82	0.0010
K10 Δ	1	79	129.83	0.10	0.76	0.0010
Edinburgh Δ	1	79	129.83	0.00	0.99	<0.001
Age	1	79	129.83	0.05	0.82	0.0010

Note: Δ = change score (PRG-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 13. Repeated Measures ANOVA results for PRE and PRG timepoints across all three groups for Paired Recognition memory performance.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	152.67	1.61	0.21	0.038
Session	1	81	26.63	13.18	< 0.001	0.14
Group $\times$ Session	2	81	26.63	0.97	0.38	0.023
Sleep Δ	1	81	26.63	0.51	0.48	0.006
K10 Δ	1	81	26.63	0.63	0.43	0.008
Edinburgh Δ	1	81	26.63	0.44	0.51	0.005
Age	1	81	26.63	3.48	0.066	0.041

Note: Δ = change score (PRG-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .017.

Supplementary Table 14. Repeated Measures ANCOVA results for PRE and POST timepoints across all three groups for MASQ Language subscale.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	39.86	0.20	0.82	0.0050
Session	1	81	13.59	46.99	< 0.001	0.37
Group $\times$ Session	2	81	13.59	2.94	0.058	0.068
Sleep Δ	1	81	39.86	0.77	0.38	0.0090
K10 Δ	1	81	39.86	2.52	0.12	0.030
Edinburgh Δ	1	81	39.86	0.27	0.61	0.0030
Age	1	81	39.86	0.63	0.43	0.008

Note: Δ = change score (POST-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .025.

1 Supplementary Table 15. Repeated Measures ANCOVA results for PRE and POST
2 timepoints across all three groups for MASQ Verbal Memory subscale.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	81	48.46	0.40	0.67	0.010
Session	1	81	15.81	53.42	< 0.001	0.397
Group $\times$ Session	2	81	15.81	4.83	0.010	0.107
Sleep Δ	1	81	48.46	0.20	0.65	0.0030
K10 Δ	1	81	48.46	0.37	0.54	0.0050
Edinburgh Δ	1	81	48.46	0.06	0.81	0.0010
Age	1	81	48.46	2.79	0.099	0.033

3 *Note: Δ = change score (POST-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α =*
4 *.025.*

5

Supplementary Table 16. Post-hoc for repeated Measures ANCOVA results for PRE and POST timepoints across all three groups for MASQ Verbal Memory subscale.

Group	Session	Est.	SE	df	t	p _{all}
PRG ₁	PRE - POST	-6.06	1.07	81	-5.65	<.001
PRG ₂	PRE - POST	-5.40	1.19	81	-4.54	<.001
CTR	PRE - POST	-1.74	1.03	81	-1.70	0.093

Note: p = Tukey HSD adjusted p -value. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy, POST = early postpartum.

Supplementary Table 17. Repeated Measures ANCOVA results for PRE and PRG timepoints across all three groups for MASQ Language subscale.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	84	33.48	0.04	0.96	0.0010
Session	1	84	9.54	64.30	< 0.001	0.43
Group $\times$ Session	2	84	9.54	8.53	< 0.001	0.17
Sleep Δ	1	84	33.48	0.59	0.44	0.0070
K10 Δ	1	84	33.48	0.06	0.80	0.0010
Edinburgh Δ	1	84	33.48	4.15	0.045	0.047
Age	1	84	33.48	0.55	0.46	0.0070

Note: Δ = change score (PRG-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .025.

Supplementary Table 18. Post-hoc for repeated Measures ANCOVA results for PRE and PRG timepoints across all three groups for MASQ Language subscale.

Group	Session	Est.	SE	df	p	t
PRG ₁	PRE - PRG	-6.08	0.77	84	< 0.001	-6.08
PRG ₂	PRE - PRG	-3.55	0.93	84	< 0.001	-3.55
CTR	PRE - PRG	-1.54	0.80	84	0.056	-1.54

Note: p = Tukey HSD adjusted p -value. PRE = before pregnancy; POST = early postpartum. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls.

Supplementary Table 19. Repeated Measures ANCOVA results for PRE and PRG timepoints across all three groups for MASQ Verbal Memory subscale.

Effect	df ₁	df ₂	MSE	F	p	η^2_p
Group	2	84	39.49	0.26	0.77	0.006
Session	1	84	10.85	69.89	< 0.001	0.45
Group $\times$ Session	2	84	10.85	8.27	< 0.001	0.17
Sleep Δ	1	84	39.49	0.23	0.630	0.0030
K10 Δ	1	84	39.49	0.25	0.621	0.0030
Edinburgh Δ	1	84	39.49	4.18	0.044	0.047
Age	1	84	39.49	7.58	0.0070	0.083

Note: Δ = change score (PRG-PRE). η^2_p = partial eta squared. Bonferroni-corrected threshold for Group $\times$ Session: α = .025.

1 Supplementary Table 20. Post-hoc for repeated Measures ANCOVA results for PRE and
 2 PRG timepoints across all three groups for MASQ Verbal Memory subscale.

Group	Session	Est.	SE	df	p	t
PRG ₁	PRE - PRG	-5.48	0.82	84	< 0.001	-5.48
PRG ₂	PRE - PRG	-5.64	0.99	84	< 0.001	-5.64
CTR	PRE - PRG	-1.26	0.85	84	0.14	-1.26

3 *Note: p = Tukey HSD adjusted p-value.*

4

1 Supplementary Table 21. Post-hoc results for MASQ Language subscale in PRG₁ for the
 2 PRE, POST, and POST+1 timepoint.

Contrast	Est.	SE	df	t	p
PRE - POST	-4.29	1.18	23	-3.63	0.0039
PRE - POST1	-3.83	1.26	23	-3.037	0.015
POST - POST1	0.46	1.11	23	0.41	0.91

3 *Note: PRE = before pregnancy; POST =early postpartum, POST+1 = late postpartum. P = Tukey HSD*
 4 *adjusted*

1 Supplementary Table 22. Post-hoc results for MASQ Verbal Memory subscale in PRG₁ for
 2 the PRE, POST, and POST+1 timepoint.

Contrast	Est.	SE	df	t	p
PRE - POST	-4.79	1.27	23	-3.76	0.0028
PRE - POST1	-3.21	1.17	23	-2.75	0.029
POST - POST1	1.58	1.35	23	1.18	0.48

3 *Note: PRE = before pregnancy; POST =early postpartum, POST+1 = late postpartum. P = Tukey HSD*
 4 *adjusted*

1 Supplementary Table 23. Correlations between objective memory performance and
2 subjective memory complaint difference scores (POST - PRE) per group.

Group	Objective measure	Subjective measure	N	Method	r	p	p _{FDR}
PRG ₁	Cued Recall	MASQ Language	34	Pearson	-0.19	0.28	0.42
PRG ₁	Cued Recall	MASQ Verbal Mem	34	Pearson	-0.42	0.012	0.058
PRG ₁	Item Recog	MASQ Language	34	Pearson	-0.10	0.58	0.74
PRG ₁	Item Recog	MASQ Verbal Mem	34	Pearson	-0.44	0.010	0.058
PRG ₁	Paired Recog	MASQ Language	34	Pearson	0.27	0.12	0.21
PRG ₁	Paired Recog	MASQ Verbal Mem	34	Pearson	0.08	0.65	0.78
PRG ₂	Cued Recall	MASQ Language	26	Pearson	-0.23	0.26	0.42
PRG ₂	Cued Recall	MASQ Verbal Mem	26	Pearson	-0.34	0.085	0.19
PRG ₂	Item Recog	MASQ Language	26	Pearson	-0.40	0.041	0.15
PRG ₂	Item Recog	MASQ Verbal Mem	26	Pearson	-0.38	0.057	0.17
PRG ₂	Paired Recog	MASQ Language	26	Pearson	-0.36	0.070	0.18
PRG ₂	Paired Recog	MASQ Verbal Mem	26	Pearson	-0.32	0.11	0.21
CTR	Cued Recall	MASQ Language	33	Spearman	-0.14	0.45	0.63
CTR	Cued Recall	MASQ Verbal Mem	33	Spearman	-0.04	0.82	0.84
CTR	Item Recog	MASQ Language	33	Spearman	-0.07	0.72	0.81
CTR	Item Recog	MASQ Verbal Mem	33	Spearman	0.04	0.84	0.84
CTR	Paired Recog	MASQ Language	33	Spearman	-0.52	0.002	0.034*
CTR	Paired Recog	MASQ Verbal Mem	33	Pearson	-0.43	0.013	0.058

1 PRG_1 = first-time mothers; PRG_2 = second-time mothers; CTR = nulliparous controls.

2

Supplementary Table 24. Means and standard deviations for the sleep, psychological distress, and depression scores for the main cohort.

Behavioral measure	Group	PRE	PRG	POST	POST+1
Sleep (hours a night)	PRG ₁	7.74 ± 1.46	8.09 ± 0.67	7.08 ± 1.43	7.14 ± 0.08
	CTR	7.42 ± 1.06	7.62 ± 0.88	7.60 ± 1.15	-
	PRG ₂	7.21 ± 1.77	7.68 ± 0.67	6.64 ± 1.3	-
Stress (K10)	PRG ₁	6.90 ± 3.4	7.63 ± 4.47	7.21 ± 6.14	6.40 ± 5.58
	CTR	8.6 ± 5.45	8.85 ± 5.69	9.82 ± 6.69	-
	PRG ₂	5.59 ± 4.82	9.07 ± 6.89	6.68 ± 6.12	-
Depressive scores (EPDS)	PRG ₁	11.22 ±	11.17 ±	11.21 ±	11.54 ±
		2.30	2.05	2.25	2.46
	CTR	11.55 ±	12.21 ±	12.07 ±	-
		2.18	3.30	3.06	
	PRG ₂	11.31 ±	12.29 ±	11.71 ±	-
		1.95	2.48	2.07	

Note: PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy; PRG = during pregnancy, POST = early postpartum, and POST+1 = late postpartum

1 Supplementary Table 25. Means and standard deviations for the sleep, perceived stress, and
2 depression scores for the replication cohort.

Behavioral measure	Group	PRE	PRG	POST
Sleep (hours a night)	PRG ₁	8.30 ± 8.38	7.13 ± 1.26	6.55 ± 1.13
	CTR	7.32 ± 1.12	7.49 ± 0.91	7.58 ± 1.01
Stress (PSS)	PRG ₁	14.32 ± 4.67	14.20 ± 5.51	14.69 ± 5.39
	CTR	14.90 ± 5.98	13.65 ± 6.15	14.72 ± 6.63
Depressive scores (EPDS)	PRG ₁	6.03 ± 3.45	5.71 ± 3.44	6.81 ± 3.80
	CTR	6.66 ± 4.48	6.26 ± 4.64	6.80 ± 5.13

3
4
5

Supplementary Table 26. Repeated Measures ANOVA result for PRE and POST timepoint across all three groups for hours of sleep the past week.

Analysis	Effect/Comparison	df	F/t	p-value	Mean Diff	95% CI
<i>Main ANOVA</i>						
Between-subjects	Group	2	2.35	0.10	-	-
Within-subjects	Timepoint	1	3.94	0.050	-	-
Interaction	Group × Timepoint	2	3.25	0.044*	-	-
<i>Between-groups (PRE)</i>						
	Group effect	2	1.52	0.22	-	-
<i>Between-groups (POST)</i>						
	Group effect	2	4.00	0.022*	-	-
	CTR vs PRG ₂	-	-	0.018*	0.96	[0.14, 1.78]
	CTR vs PRG ₁	-	-	0.20	0.58	[-0.21, 1.37]
	PRG ₂ vs PRG ₁	-	-	0.52	-0.38	[-1.21, 0.44]
<i>Within-groups (PRE vs POST)</i>						
	PRG ₁	30	2.58	0.015*	-0.76	[0.16, 1.36]
	PRG ₂	26	1.35	0.19	-0.48	[-0.25, 1.21]
	CTR	31	-1.03	0.31	0.22	[-0.65, 0.21]

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Mean Diff = mean difference in hours of sleep the past week. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls. PRE = before pregnancy; POST = after pregnancy.

Supplementary Table 27. Repeated Measures ANOVA result for PRE and POST timepoint across all three groups (PRG₁ n = 39; PRG₂ = 28; CTR = 40) for level of stress measured with K10 questionnaire.

Analysis	Effect/Comparison	df	F/t	p-value	Mean Diff	95% CI
Between-subjects	Group	2	3.87	0.024	-	-
Within-subjects	Timepoint	1	3.09	0.082	-	-
Interaction	Group × Timepoint	2	0.30	0.74	-	-

Note: **p* < 0.05, ***p* < 0.01, ****p* < 0.001. Mean Diff = mean difference in level of stress, higher scores reflect higher self-reported levels of stress. PRG₁ = first-time mothers; PRG₂ = second-time mothers; CTR = nulliparous controls.

1 Supplementary Table 28. Repeated Measures ANOVA result for PRE and POST timepoint
 2 across all three groups (PRG₁ n = 39; PRG₂ = 28; CTR = 40) for depressive scores measured.

Analysis	Effect/Comparison	df	F/t	p-value
Between-subjects	Group	2	1.03	0.36
Within-subjects	Timepoint	1	1.89	0.17
Interaction	Group × Timepoint	2	0.40	0.68

3