Development of the Paternal Brain in Humans throughout Pregnancy

Françoise Diaz-Rojas (fdiazro01@gmail.com)
Kyoto University https://orcid.org/0000-0003-1989-6049
Michiko Matsunaga
Kyoto University
Yukari Tanaka
Kyoto University
Takefumi Kikusui
Department of Animal Science and Biotechnology, Azabu University
Kazutaka Mogi
Azabu University
Miho Nagasawa
Azabu University
Kohei Asano
Osaka University of Comprehensive Children Education https://orcid.org/0000-0001-8902-3804
Nobuhito Abe
Kyoto University https://orcid.org/0000-0002-9897-4414
Masako Myowa
Kyoto University https://orcid.org/0000-0001-6080-106X

Article

Keywords: pregnancy, paternal, neural, postpartum, caregiving

Posted Date: November 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-934033/v2

Abstract

Previous studies have demonstrated that paternal caregiving behaviours are reliant on neural pathways similar to those supporting maternal care. Interestingly, a greater variability exists in parental phenotypes in males than in females among individuals and mammalian species. However, less is known about when or how such variability emerges in human males. We investigated the longitudinal changes in the neural, hormonal and psychological bases of expression of paternal caregiving in humans throughout pregnancy and the first four months postpartum. We measured oxytocin and testosterone, paternity-related psychological traits and neural response to infant-interaction videos using fMRI in first-time fathers and childless men at three time points (early–mid-pregnancy, late-pregnancy and postpartum). We found that paternal-specific brain activity in prefrontal areas distinctly develops during middle-to-late pregnancy and is enhanced postpartum. Additionally, among fathers, the timing of the development of prefrontal brain activity was associated with specific parenting phenotypes.

Introduction

Humans are biparental mammals, that is, males and females work together to care for their offspring. Many studies have demonstrated that the increased involvement of fathers in childrearing augments the mental health and social adaptation of the child and mother[1–4]. However, beyond maternal care, paternal caregiving shows considerable variability throughout history and across cultures. Many fathers directly engage in childrearing activities, whereas others hardly do so[5–8]. Thus, understanding the factors that modulate such variability in parental phenotypes in men is of great importance given that paternal caregiving influences the development of the offspring[9].

The fact that human paternal behaviour is highly variable attracted a strong scholarly interest in its neural and hormonal substrates with several studies on animal models[10,11] and fMRI studies on humans. With regard to studies on humans, variations in paternal phenotypes have been considered to be associated with differences in the neural activity of several areas in the brain[12]. These areas together form the parental brain networks that are associated with parental caregiving behaviour in humans[8,13]. These areas consist of the mentalising[14,15], the emotional processing[16,17], the embodied simulation[18,19] and the subcortical mammalian parenting networks[8,15].

Importantly, the parental brain is not exclusive to parents. Both non-fathers and fathers display neural activation in these networks in response to infant-interaction stimuli[20,21]. However, the fact that differences exist in the patterns of activation between the two groups is more important. Among fathers, the patterns of activation are influenced by active caregiving experiences; thus, primary caregiving fathers (i.e. stay-at-home fathers) exhibited increased activation of the emotional processing network compared with that of secondary caregiving fathers[12]. In addition, from an anatomical standpoint, Kim et al. (2014)[22] found that a postpartum decrease in grey matter volume in the orbitofrontal cortex was associated with high levels of father–infant interaction among first-time fathers. Thus, plastic changes at

the functional and structural levels in the human parental brain are related to caregiving experiences with infants.

However, less is known about when such paternal phenotypical and biological changes occur. Thus far, very few studies have focused on the development of the human paternal brain from early pregnancy to postpartum[13]. Recently, Diaz-Rojas et al. (2021)[21] found changes in the parental brain from early to mid-pregnancy (gestational age [GA] < 30 weeks). Using multivoxel pattern analysis (MVPA), they identified subtle differences in areas of the parental brain between non-fathers and first-time fathers in response to infant-interaction stimuli. In particular, differences were found in the mentalising areas between fathers with and without previous experience of infant caregiving (i.e. infants apart from their own). Considering these findings, it is possible that the variability in parental brain and phenotypes begin forming during the pregnancy period, long before the father acquires active caregiving experiences with his child.

In addition to active caregiving experiences, other factors may modulate such differences. For example, although males are not exposed to radical biological changes compared with those of females during pregnancy and postpartum, evidence exists that their hormonal profiles change during this period[23–25]. Two of these hormones are oxytocin and testosterone, which play an important role in paternal behaviour[26,27]. Low levels of testosterone throughout pregnancy and during postpartum in fathers were associated with high levels of paternal involvement in caregiving[27,28]. Moreover, high levels of postpartum oxytocin were related to increased father–infant interaction[29]. In this manner, these hormonal changes may modulate the variability in parental phenotypes among first-time fathers even during pregnancy.

To clarify the dynamism of paternal brain development and its relation to parenting behavior and hormones, continually examining the evolution of the parental brain is necessary from the early period of the partner's pregnancy until postpartum. In this regard, the current study investigated longitudinal changes in the neural, hormonal and psychological bases of the expression of human paternal caregiving across pregnancy and after childbirth. Using fMRI, we recorded neural activation in the parental brain in response to infant-interaction stimuli among first-time fathers. The study employed three time points, namely, early–mid-pregnancy (GA: 10–30 weeks), late-pregnancy (GA: >30 weeks) and in the first four months postpartum. We also examined the abovementioned changes among childless males as the control group during the same time periods to serve as a baseline. Moreover, we observed the levels of two parenting-related hormones, namely, oxytocin and testosterone, and paternity-related psychological and behavioural traits for each time point. We first investigated changes in terms of group differences (first-time fathers and childless males) to reveal whether and how the neural networks of the parenting brain of first-time fathers change during pregnancy. Next, we compared the hormonal, psychological and behavioural profiles of first-time fathers to investigate which of these factors are related to the development of the paternal brain. Finally, we compared differences in the timing of neural changes in expectant fathers to examine the possible factors that influence the paternal brain from various perspectives.

We demonstrate that across pregnancy, first-time fathers experience a major change in their neural response to infant-interaction stimuli, especially in the mentalising network compared with control across the same time periods. Moreover, activations in several areas of the paternal brain among postpartum but not prepartum fathers, are related to emotional attachment to their infants. Finally, the timing of changes in the mentalising network is related to several phenotypes, such as a father's feeling of attachment towards his own infant and positive outlook towards parenting.

Results

We recruited 72 men: 36 men without children as the control group (CG) and 36 first-time expectant fathers in the early–mid-pregnancy (GA: 10–30 weeks) of their partners for the 'papa' group (PG). Each subject participated in an fMRI scanning session (session 1—early-pregnancy for PG) to record neural activation in response to videos that display infant–adult interaction scenarios. The stimulus videos consisted of four scenes, namely, the first two showing male–infant social interactions (S1: playing with an infant; S2: changing diapers) and the second two showing control non-infant interactions (C1: opening a box and removing a tripod; C2: wrapping a box with plastic). All videos were shown from the first-person view (Fig 1). The motor movements of each control scene roughly matched those of the corresponding infant-interaction scene. Approximately four months later, the participants were invited again for a second fMRI measurement (session 2—late-pregnancy for PG), and for an additional time subsequently four months later (session 3—postpartum for PG). At each session, we measured various behavioural traits related to parenthood, relationship with the partner, socioeconomic status (SES) and others (see the Methods section for a complete description of all recorded traits). Saliva samples were obtained to analyse oxytocin and testosterone. Table 1 summarises these data.

Parental brain activities at the three time points

We were interested in the cross-sectional snapshots of the paternal brain at each time point (i.e. early and late-pregnancy and postpartum) to reveal whether and the extent to which the brain response to infant stimuli of fathers would differ from those of non-fathers. Moreover, we aimed to determine changes in the neural activities in terms of longitudinal developmental processes from pregnancy to the postpartum period.

First, we examined the paternal brain of expectant fathers and childless males using a whole-brain analysis at each time point. Diaz-Rojas et al. (2021)[21] reported the cross-sectional results of the early-pregnancy period (session 1, GA < 30 weeks); therefore, we excluded them from the current study. For session 2 (late-pregnancy), we examined the overall response to the infant stimuli versus the control stimuli (i.e. combination of S1 and S2 vs combination of C1 and C2, or the S–C contrast) in expectant fathers and control men to confirm the activation of the paternal brain at this stage. We found that, when taken together, all participants displayed widespread activation across the brain (Fig 2a, see also Table

S1). These activated areas correspond mainly to the human parental caregiving networks[8,12], that is, an *emotional processing network* (i.e. inferior frontal gyrus, anterior cingulate cortex, insula and amygdala) and a *mentalising network* (i.e. the medial prefrontal cortex (PFC), temporal poles and superior temporal sulcus). A comparison between the average brain activities of the PG and the CG revealed decreased activation in the right dorsomedial prefrontal cortex (dmPFC) in PG than in CG (Table 2). Regarding session 3, we found similar activations in the same areas of the paternal brain (Fig 2b and Table S2), although they involve a larger portion of the brain. In addition, the PG exhibited increased activation in the left dmPFC at postpartum in comparison with the CG (Table 2).

In summary, these results confirm the findings of Diaz-Rojas et al. (2021), that is, the parental brain is a staple of the male brain regardless of paternity status. Conversely, we observed substantial differences between the neural activities of fathers and non-fathers in response to the infant caregiving context, especially in the mentalising areas of the human caregiving network.

Second, we focused on the relationships between the cross-sectional brain response toward infant stimuli and behavioural or hormonal traits related to parenting. We examined the relationships in both groups of participants in session 2 and session 3 and used whole-brain analysis with individual regression models for each behavioural and hormonal covariate and the S–C contrast. Table 2 summarises the clusters in the brain, which exhibited significant correlations with behavioural and hormonal covariates. In session 3 (postnatal period), we noted several clusters of brain activity correlated to several behavioural and hormonal traits in the PG. Of interest was *postnatal attachment*, which indicated a strong and widespread negative relationship with the activation from the S–C contrast, mainly across the emotional network (insula, ACC, dorsolateral PFC and supramarginal gyrus) and to a lesser degree with the mentalising network (precuneus). Alternatively, we found neither *foetal attachment* (the prenatal counterpart to *postnatal attachment*) nor any other behavioural traits correlated with the S–C contrast in session 2 (late-pregnancy period) except for household income in CG. For the hormonal measurements, we found only a negative relationship between oxytocin and the right supplementary motor cortex in PG in session 2. No other brain areas in session 2 or 3 showed any other relationship with oxytocin or testosterone in the PG and CG.

Longitudinal development of the paternal brain

The question remains about whether and how expectant fathers' neural activation patterns observed in response to infant stimuli begin to change across the pregnancy period to childbirth. Continuing the whole-brain approach, we compared changes in the neural response to the infant-interaction stimuli in both groups from one time point to another and between groups. For the PG, we detected a change from session 2 to session 3 in the left dmPFC (Table 3). Conversely, we were unable to detect any differences in the PG between pregnancy sessions (from session 1 to 2), or in CG between any of the time points. Comparing the two groups revealed an interaction effect between session and group in the

dmPFC (left and right). A further exploration of this interaction revealed that the dmPFC in PG displayed an increased response from the session 1 to session 3 than that of the CG. The PG exhibited no major difference in responses from session 1 to 2, thus, we averaged them as the *pregnancy session* to view the overall change in the paternal brain from pregnancy to postpartum. We found that fathers remarkably exhibited increased activation in the dmPFC from pregnancy to the postpartum period compared to brain activities observed in the CG across the same time period (Fig 3A).

MVPA for paternal brain development

These results suggest that changes in the neural response to infant-interaction visual stimuli in the dmPFC in expectant fathers occur around the delivery period without an apparent change during pregnancy. However, we need to consider the possibility that the whole-brain analysis may lack sufficient statistical power to detect smaller changes in the brain of expectant fathers during this period. Actually, Diaz-Rojas et al. (2021) found that small changes in the development of the paternal brain in early pregnancy (~20 weeks GA) could be evidenced using multivoxel pattern analysis (MVPA). Therefore, we used MVPA to examine the multivoxel patterns of activation of the dmPFC across the two time points during pregnancy.

To conduct MVPA, we used a support vector machine (SVM) model that performs a supervised classification of session (early- or late-pregnancy) using the multivoxel data from the left dmPFC, as it was the area that consistently displayed the majority of changes from pregnancy to postpartum. In the PG, the algorithm successfully classified the pregnancy period with a higher-than-chance accuracy (mean ± 95% confidence interval [CI]: 66.17% ± 13.25, $p = 0.005$, true negative rate [TNR] = 0.66, true positive rate [TPR] = 0.67, Fig 3B). In other words, the multivoxel activity patterns of the left dmPFC differ between early and late-pregnancy. Conversely, the algorithm could not classify the session in the CG with a higher-than-chance accuracy (55.88% ± 13.91, $p = 0.19$, TNR = 0.54, TPR = 0.58). This finding suggests that the underlying neural encoding patterns in non-fathers remain the same between the two time points (session 1 and 2).

In addition, we classified group (PG or CG) in each session (1, 2, and 3) using MVPA to determine whether the activity patterns of the left dmPFC during pregnancy were specific to fathers. We found that the classification of paternal status (i.e. belonging to the PG or CG) was possible using the multivoxel data of this ROI at either session 2 and 3 respectively (session 2, PG vs CG: 64.70% ± 11.65, $p = 0.01$, TNR = 0.66, TPR = 0.64; session 3: 65.67% ± 11.67, $p = 0.007$, TNR = 0.63, TPR = 0.66) but not from the data of session 1 (session 1, PG vs CG: 55.88% ± 12.11). These results suggest that the activity patterns of the dmPFC in response to infant-related stimuli will remarkably change only among expectant fathers, especially after mid or late-pregnancy (Fig 3C).

Effect of the dynamics of the paternal brain development on parental phenotypes

Our findings suggest that the paternal brain, especially in the medial prefrontal cortex, in expectant fathers, mainly develops across the last weeks of pregnancy to the early period of childbirth. However, previous studies reported that various types of change occur in the developmental processes of paternity. Berg et al. (2001)[23] proposed that men display distinct patterns of variations in testosterone across the pregnancy period. Moreover, Diaz-Rojas et al. (2021) found differences in neural encoding patterns between expectant fathers during early pregnancy. After childbirth, Abraham et al., (2014)[12] suggested that not all men exhibited the same patterns of development in the paternal brain. Given these findings, the current study conducted *post hoc* exploratory analysis to further investigate individual differences in the development of the paternal brain in postnatal fathers and explored the factors that may contribute to it. We focused on the individual differences observed with dmPFC activation, which indicated larger changes from pregnancy to postpartum. In addition, we examined the relationship of the degree of change in activation in response to infant-interaction stimuli with psychological/behavioural characteristics and hormone concentrations in the PG.

One method for quantifying these developmental profiles is to categorise subjects on the basis of their progression in dmPFC activation across the pregnancy. However, using this method may yield an excessive number of groups given the limited sample of fathers. Thus, we performed automatic partitioning of the sample into two groups using *k*-means clustering based on changes in the left dmPFC in response to infant stimuli from session 1 to 2, and from session 2 to 3 instead of manually separating fathers into groups. Consequently, data were separated into two sub-groups, namely, group 1 (exhibited relatively high levels of change from session 1 to 2, but low levels of change from session 2 to 3 (n = 20) and group 2 (displayed relatively low levels of change from session 1 to 2, but high levels of change from session 2 to 3 (n = 12; Fig 4A and B). Using a two-way ANOVA, we found an overall session effect for both groups ($F = 5.60$, $p = 0.005$, d.f. = 2), but not a group ($F = 1.20$, $p = 0.27$, d.f. = 1) nor an interaction effect ($F = 2.88$, $p = 0.061$, d.f. = 2).

Using these categories, we compared hormonal profiles and behavioural scores across pregnancy among the sub-groups and observed important differences (Fig 4C; Table S3 for all statistical tests). Specifically, *positive image towards parenting* and *foetal/postnatal infant attachment* pointed to differences between the sub-groups ($F = 11.68$, $p < 0.0001$, and $F = 8.13$, $p = 0.0002$, respectively). *Weekly worktime* also showed marginal group differences ($F = 4.68$, $p = 0.038$), but after correction for multiple comparisons, this difference did not reach a significant level. Despite these behavioural differences, both sub-groups had similar average levels of dmPFC activation at session 3 (Fig 4B), which suggests that the developmental trajectories that each group undergoes during pregnancy may influence the type of father each man will become.

Discussion

This study aimed to elucidate when and how the paternal brain develops in first-time fathers from the early-pregnancy period to four months postpartum. We found that parts of the mentalising network among first-time fathers remarkably change from mid-pregnancy (GA > 30 weeks) to the postpartum period.

Particularly, the study obtained four findings. First, differences are evident in the activation in the dmPFC between prenatal and postnatal recordings in first-time fathers but not in control men between the same time points. Second, using MVPA, we determined different neural patterns of activation in the left dmPFC between (a) early–mid-pregnancy and late-pregnancy recordings in fathers and (b) between late-pregnancy fathers and the control group. In other words, the activity patterns of expectant fathers in the left dmPFC for infant-related stimuli substantially changed from mid-pregnancy onwards. Third, a strong relationship exists between emotional attachment towards their infants and neural activation in response to infant-interaction stimuli among postnatal but not prenatal fathers. Finally, differential developmental trajectories in the dmPFC were related to different paternal phenotypes.

Regarding the first finding, fathers (after childbirth) displayed an increased activation in the dmPFC compared with their recordings during pregnancy (before childbirth) and that of childless control males. The dmPFC is one of the main parts of the mentalising network[15]. In infant caregiving contexts, the mentalising network is considered to be involved in imaging or reasoning of the psychological states of infants through nonverbal signals, such as facial expressions[8]. Such active caregiving experiences with their infants in everyday life may enhance the dmPFC activation of fathers after childbirth. In fact, Abraham et al. (2014)[12] showed that the time fathers spent in childcare is correlated with connectivity between the amygdala and the superior temporal sulcus, which are parts of the parenting brain networks. In addition, Diaz-Rojas et al. (2021)[21] revealed that past experience with infant caregiving (e.g. taking care of a younger sibling/nephew) modulated the activity in certain areas of the mentalising network in expectant fathers. Thus, active caregiving experience may be a strong modulator of the development of the paternal brain in humans.

Importantly, the current study provides evidence that the development of the parental brain does not start after childbirth. Instead, a certain degree of change in the dmPFC occurs from the early to the late-pregnancy period. Thus, MVPA could identify significant changes between early-mid-pregnancy and late-pregnancy fathers using the multivoxel neural patterns of the dmPFC. This finding suggests a notable difference between the neural encoding mechanisms between the two time points. Furthermore, we discriminated the difference between late-pregnancy fathers and control, but, importantly, not between early-mid-pregnancy fathers and their control. These findings support the assumption that although direct caregiving experiences with their infant may strengthen the parental brain, a certain of degree of development is initiated during the pregnancy period before the actual caregiving contexts.

Interestingly, no other brain region showed similarly strong differences in activation in response to infant-interaction stimuli throughout pregnancy compared with the dmPFC. In the same manner, the study observed the lack of widespread differences in the neural response between the PG and CG. For first-time

fathers and childless males, the neural response to the stimuli in other areas of the parental brain remains relatively constant across the three sessions. One possible reason is that expectant fathers and childless males display a well-defined activation pattern in the parental brain when exposed to infant stimuli[21] (Fig 2A). Thus, the strengthening of the paternal brain response after childbirth may effectively be limited due to a ceiling effect. Another possibility is that further development of the paternal brain is dependent on experiences of infant-interaction, as argued by previous studies. Primary caregivers, regardless of gender, were found to exhibit a stronger activation of the emotional processing network when watching videos of their infant compared with fathers as secondary caregiver[12]. According to this, it may be interpreted that the fathers participating in this study were still lacking in caregiving experiences necessary for the development of other areas of the parental brains, particularly those related to emotional processing. Indeed, on average, the participants in the father group had their postnatal recording session (session 3) at approximately four months after childbirth. At this period, the infant mainly relies on maternal care (e.g. breast feeding) through skin-to-skin contact with the mother. Consequently, opportunities for direct father–infant interactions may be limited. Furthermore, the majority of the PG participants (77%) reported spending 7 h per week or less alone with their infant; a third of the PG participants indicated spending less than 1 h per week. We were unable to establish a direct link between time spent with their infant and the development of the paternal brain. Thus, the current data may reflect only the first stage in developmental processes of the postnatal paternal brain.

In terms of the third finding, *postnatal infant attachment* exerted a negative linear relationship with the neural response to infant stimuli in several brain areas, including the mentalising (precuneus) and emotional processing networks (left insula, anterior cingulate and the left lateral PFC). How, then, should these results be interpreted in terms of postpartum fathers? One possibility is that fathers with increased attachment towards their infant may feel a stronger disconnect with unfamiliar infants presented in the stimuli, which led to the reduced neural response. Previous studies indicated that parents' preference for their infants in comparison with unknown infants is reflected at the neural level[30–33]. Thus, the emotional processing network is considered to be involved in emotional and cognitive empathy, including the understanding of an infant's emotions and pain[34,35]. Particularly, the left insula has been shown to have a greater activation in fathers in response to images of their own children versus those of other children[32]. Notably, the same relationship was not observed between prenatal foetal attachment and the neural response to infant-interaction stimuli (before childbirth). A systematic review of studies on prenatal and postnatal parent–infant attachment suggested that little conclusive evidence links prenatal foetal attachment to postnatal parent–infant relationship in fathers[36]. Although the current study could not elucidate the actual psychological mechanism that underlies the feeling of attachment of fathers towards their infants, the link between the paternal brain and feeling of attachment may possibly only form at postpartum following real caregiving experiences.

Finally, the timing of the changes in the response of the dmPFC was related to particular phenotypes in expectant fathers. Fathers who showed early changes (i.e. relatively large differences between session 1 and 2) in the left dmPFC were associated with high scores for *positive attitude towards parenting* and

foetal/infant attachment than those who exhibited the change at a later time (i.e. between session 2 and 3). The *positive attitude towards parenting* questionnaire includes items related to emotional judgements towards the unborn infant and the act of parenting, as well as ratings of personal growth and satisfaction towards parenting[37]. *Postnatal attachment* was correlated with the score for *father development* (Fig S1), which is a metric associated the fathers' feeling of developing a sense of fatherhood and personal growth. On the prenatal side, high levels of foetal attachment were related to a more balanced representation of the unborn infant[38] among first-time fathers. These metrics reflect mental processes modulated by the dmPFC, which is involved in understanding one's identity[39], self- and other-judgements[40] and emotion regulation[41]. Expectant fathers with early changes in the dmPFC may gain a head start compared with fathers with late dmPFC changes. We also found a difference in average weekly working hours between Group 1 and Group 2 fathers, albeit only marginal (i.e., before correction for multiple comparisons). In previous studies, working hours have been considered to be a factor that affects the expression of paternal nurturing behaviour[42,43]. It also may be related to individual differences in the dmPFC activity and its developmental trajectories shown in the present study; however, further research is needed to confirm this. It is important to highlight that the majority of the participants had high scores across the behavioural metrics related to parenting. In comparison with other studies, our sample may be positively biased and may show the differences that exists in the high-end of the paternal spectrum or simply may be insufficiently varied to provide an appropriate comparison with previous findings.

Scientists propose that testosterone and oxytocin play an important role in the parenting behaviour of males[24,27]. The current study found a transient negative correlation between oxytocin and the supplementary motor area during late-pregnancy in fathers. The supplementary motor area is involved, among other things, in motor imagery[44]. In caregiving contexts, a recent review of fMRI studies of the paternal brain[13] summarised that the supplementary motor area also displays a response to infant stimuli. Thus, the area was proposed to be involved in the creation of time-dependent motor memories, which are crucial for the development of the imagery of self–other interactions, including parent–infant exchanges[45]. People with high oxytocin levels may easily imagine these exchanges when observing the father–infant interaction videos, especially in comparison with control, which leads to the reduced activation. Alternatively, we were unable to identify differences in testosterone levels between the time points, or to determine any correlation between testosterone level and neural response to the infant stimuli. Berg and Wynne-Edwards (2001)[23] demonstrated a sharp decline in testosterone levels among expectant fathers in the period around the postpartum. Such a decrease in testosterone level is seemingly time-sensitive, which is limited to a one-month time window after childbirth[46]. Based on these findings, measuring testosterone levels earlier in the postpartum period may be necessary. Another possibility for the lack of findings may be related to the large variability in hormonal measurements, particularly given the limited sample size. Thus, further studies are required to consider the timing of hormonal measurements and the large degree of variability to properly assess their impact on the development of the parental brain.

One of the concerns arising from our findings is the ambiguity related to whether the neural activations observed in response to infant stimuli reflect a so-called “parental brain”, or more general “social brain”. Due to the characteristics of the stimuli used to investigate the parenting brain (e.g., Abraham et al. (2014)[12], but see also the Supplementary Discussion for further discussion on the stimuli), it is difficult to discriminate a child-related component (i.e., parenting) from a general social component in the brain activations in response to our stimuli. A recent systematic review of the fMRI studies in human fathers[13] has noted essential difficulties in measuring the parental brain networks solely, since they overlap with other social networks. Moreover, the issue of whether the parental brain would be a separate, specialized network from those of the social brain is one that, to our knowledge, has not been solved. In light of this, the label of “parental brain” used in this study follows the conventions of similar other studies[8,13], and refers to the areas that respond to infant stimuli. Most noteworthy is that we could identify father-specific differences in some of these brain areas in contrast with those of the control men. Even if the response to the infant stimuli is overlapped with those to social ones, the findings clearly show that these were based on parental-specific experiences.

In conclusion, we found strong evidence that the development of the paternal brain begins from the pregnancy period, a time when expectant fathers lack experience in active caregiving with their child. Importantly, we also presented evidence that the timing of the development of the paternal brain is related to different parenting phenotypes. These findings may serve as reference for the formulation of parental support programmes for expectant fathers (i.e. during pregnancy), which can be applied to diverse individuals. This direction may prove particularly effective in males who struggle to establish their sense of fatherhood and who may be at risk of committing infant abuse.

Declarations

Data availability statement

Due to strict privacy concerns, behavioral, hormonal, and MRI data obtained throughout this study can only be made available from F.D.R. upon reasonable request. Preprocessed anonymized data (fMRI voxel data, behavioural scores and hormonal residuals) required to run the analysis are included in the study repository.

Code availability statement

All custom scripts and codes used to generate the results hereby presented in this study are available from https://github.com/fdiazro01/PapaPro2021. Main routines used for whole-brain analysis (spm12), ROI signal extraction (BrainDecoderToolbox2), and MVPA (libsvm) are available from their respective owners, as described in the Methods section.

Acknowledgments

The authors are grateful to the staff at Kokoro Research Center, Kyoto University, for making its MRI scanner and associated facilities available for the purpose of this project. This study was supported by funding from the Center of Innovation Program from Japan Science and Technology Agency, JST (JPMJCE1307) to M.Myo, and the Maekawa Foundation to Y.T. (2018-2019).

Methods

Ethical considerations

The Ethics Committee of the Kyoto University Unit for Advanced Study of Mind (30-P-8) approved the experimental protocol and procedures. All participants provided written informed consent prior to the study and were remunerated equally for their participation after each session.

Participants

We recruited 72 males from all SES groups: 36 men without children (CG: mean age = 31.1 ± 6.2 years) and 36 first-time expectant fathers (PG: mean age = 33.1 ± 5.7 years; mean GA of partner = 20.9 ± 6.2 weeks). Approximately four months later, the same men were invited for a second (late-pregnancy session) and a third (after-birth) session four months thereafter. The mean intervals (in days) between first and late-pregnancy sessions were 99.18 ± 36.24 and 125.56 ± 23.53 for PG and CG, respectively. The mean intervals between the first and after-birth sessions were 250.94 ± 38.25 and 250.00 ± 27.44 for PG and CG, respectively. In the PG, the mean interval between child birth and the third session was 119.7 ± 20.3 days. Data for one PG subject at session 1 was considered as late-pregnancy, because data were obtained when the subject's partner GA was 39 weeks (i.e. the subject was missing the data for session 1; interval between recordings = 132 days). One subject in PG did not participate in the third session (after-birth). Moreover, we excluded one PG subject from all analyses due to the lack of a late-pregnancy session and participation in the after-birth session approximately nine months after birth (250 days), which was considerably later than the rest of the fathers. In the CG, one subject was excluded from all analyses due to the inability to complete the fMRI recording session. Another subject was excluded due to only completing the first session of recording. Finally, data for one subject in the CG during the third session could not be recorded due to machine error. Moreover, we discarded the fMRI data of one PG subject (first session) and one CG subject (second session) due to high movement artefacts. The final sample is as follows: for the first session (early-pregnancy), PG = 33, CG = 34; for the second session (late-pregnancy), PG = 35, CG = 33; and for the third session (after-birth), PG = 34, CG = 33.

For the clustering analysis (see subsection "Exploratory analysis of the individual differences in the paternal brain" below), only subjects with three data points were considered. For this analysis, the subject number is 32 subjects for PG, and 32 subjects for CG.

All subjects from the PG and six from the CG were married or living with partners. Their marital status remained the same across the duration of the project. All subjects were of East Asian ethnicity (Japanese: n = 70; Korean: n = 2), had normal or corrected-to-normal vision and were all right-handed. No participants reported any major medical illnesses, major psychiatric disorders or neurological illnesses.

Behavioural characteristics

For each recording session, we collected the following demographic data: (a) socioeconomic status (SES), including household income, educational background and history and number of family members in the household and (b) weekly work/study time.

Additionally, we collected the following psychological characteristics: (c) State–Trait Anxiety Index using the State–Trait Anxiety Inventory (STAI-Trait[47]) and (d) depression index using Beck's Depression Inventory (BDI)[48]. These two psychological traits have been implicated in the heterogeneity of parenting behaviours in mothers and fathers[49], as demonstrated by a decrease in positive parenting style and an increase in negative parenting style, respectively[50].

Furthermore, we considered the following behavioural data related to family and parenting: (e) positive and (f) negative attitudes towards parenting[37], including 12 items for a positive image (e.g. parenting involves pleasure) and 15 items for a negative image (e.g. parenting is difficult). Each item was scored using a five-point Likert scale. Next, we measured (g) partner relationships using the Dyadic Adjustment Scale[51] (DAS), which was only answered by married participants and those living with romantic partners with 32 items for assessing marital relationship quality. (h) Foetal–paternal attachment (The Paternal Antenatal Attachment Scale[52] was used only for PG subjects) was used to measure the strength of attachment using 16 items. Each item was rated using a five-point Liker-type scale. The scales from (e) to (h) provided scores consistent with those of previous studies. Ceiling/floor effects were observed for several items on the (e), (f) and (h) scales, which were defined by whether the mean score of ±1 standard deviation exceeded the range of possible values. Items indicating a ceiling/floor effect were excluded, and mean scores were used for analysis. To compare the prenatal foetal–paternal attachment and postnatal infant attachment, (h) scores were normalised to the 0–1 range. For the (g) scale (DAS), two items in the scale were rated as 0 or 1, which is in contrast with the rest (i.e. using a five-, six- or seven-point Likert-type scale). Thus, these two items (i.e. 'feeling of exhaustion when thinking about intercourse with partner' and 'expressions of affection') were excluded from the overall DAS score to avoid uneven effects.

Finally, we performed additional surveys only in postpartum fathers to determine paternal behaviour. We measured fathers' development and postnatal infant attachment (normalised to the 0–1 range), paternal involvement in caregiving, time together and times at papa school. Fathers' development, which is based on Morishita[53] (2006), measures psychological changes in men through the experiences of parenting

(e.g. *considering the feelings of parents with children*) with 20-items rated using a five-point Likert-type scale. The postnatal infant attachment scale[54] was used to measure the strength of attachment with 19 items, which were rated using a five-point Likert-type scale. Paternal involvement in caregiving was developed based on Abraham et al. (2014)[12], which measures the frequency of fathers' performance of 15 daily parenting activities (e.g. bathing children, changing diapers and feeding babies). Each item was rated using a five-point Likert scale. We also recorded the time that the participants spent alone with infants in hours per week and the number of times the participants attended parent–child classes at hospitals or municipalities during their partners' pregnancy. These frequencies were respectively designated as *time together* and *times at papa school*. Moreover, we measured the degree of household chores (e.g. cooking meals, doing the laundry and cleaning the house) performed, which was developed based on Volling & Belsky (1992)[55] and Edelstein et al. (2017)[24]. Eight items were scored on five-point Likert-type scale ranging from 1 (always the wife) to 5 (always the husband).

Table 1 provides a summary of the behavioural data. Statistical differences between groups and sessions were assessed using a N-Way ANOVA. Several behavioural traits displayed group differences, such as *trait anxiety* ($F = 11.8$, $p = 0.01$, $df = 1$), *weekly worktime* ($F = 115.7$. $p < 0.001$, $df = 1$), *household income* ($F = 133.8$, $p < 0.001$, $df = 1$) and *positive attitude towards parenting* ($F = 24.7$, $p < 0.001$, $df = 1$). Despite these group differences, controlling for any of the covariates did not significantly change the differences in neural activation across the abovementioned sessions. Several traits exhibited a large correlation with one another (Fig S1). To reduce the number of tests performed across the study and to avoid confusion in our conclusions due to similarity among scores, we excluded a few of the measurements. In particular, *fathers' development* displayed a high correlation with several other parent-related traits, such as *postnatal infant attachment* and *positive attitude towards parenting* and, to a lesser extent, *time together* and *parental involvement*. For this reason, we excluded this trait from analyses, as its inclusion would not provide any useful insights for the study objectives. The depression score, *BDI*, pointed to a large correlation with *trait anxiety*. Thus, it was also excluded from the analysis. *State anxiety* was likewise excluded, because the measurement itself provides no relevant information regarding parenting. Instead, it reflects a number of factors, such as post-recording stress or anxiety and may disrupt our analysis.

Data collection and analysis for hormonal levels

The subjects rinsed their mouths prior to saliva collection and chewed on an oral cotton swab after 10 min (Salimetrics, State College, PA, USA), which was placed sublingually for 3 min. This process was repeated to obtain a backup sample. The first samples were used for the calculation of the hormonal levels of each subject. If the first sample contained insufficient saliva specimen or if the hormonal levels of the first sample were outliers, then the value of the second sample was used as the individual value. The samples were stored at −80°C until assayed. Salivary oxytocin was assayed using a commercial kit (ADI-900-153A-0001; ENZO Co. Ltd., Tokyo, Japan) following the manufacturer's protocol. In summary,

240 µL of saliva was dried using a Speedvac evaporator at room temperature for 3 h and reconstituted in 240 µL of an assay buffer out of which 100 µL was used for the assay. Salivary testosterone was measured by ELISA following our previous reports[56,57]. Furthermore, 25 µL of saliva was used in the assay using testosterone-3-CMO-HRP (FKA101; COSMO Bio, Tokyo, Japan) and a specific anti-testosterone serum (FKA102-E; COSMO Bio). All intra- and inter-assay coefficients of variation were less than 15%. We could not detect oxytocin from the samples of two subjects from the PG and two subjects from the CG. As such, these subjects were excluded from oxytocin analysis.

For the analysis of hormonal data, we addressed three nuisance effects, namely, time of day at measurement, seasonal changes and storage time. The first two refer to the fact that levels of oxytocin and testosterone vary across the day and across seasons. To minimise the first confounding effect, all experiments were scheduled at morning between 8:45 AM and 10:45 AM. However, minimising the seasonal effect at the experimental level is complex. The final nuisance effect, which may influence data quality, is the time of storage of the saliva samples before processing. Although we intended to analyse the data as soon as possible, we cannot overlook the possibility that a certain level of nuisance effect was included. Thus, to remove the three nuisance effects, we created individual linear models for each hormone using the three parameters and the raw hormonal measurement. Lastly, we obtained the residuals of the model. Across this work, the residuals were used as the values for the hormones instead of raw data.

Statistical differences between groups and sessions were assessed using a N-Way ANOVA. For PG and CG from the first session to the second and third sessions, we observed a significant decline in oxytocin levels ($F = 9.15$, $p = 0.002$) and a marginally significant increase in testosterone levels ($F = 5.21$, $p = 0.09$). However, when residualizing according to the time of day at acquisition, day of the year and time of storage before processing, these effects disappeared (session effect on testosterone: $F = 2.51$, $p = 1$) or became weaker (session effect on oxytocin: $F = 7.5$, $p = 0.01$).

Task protocol

We followed the protocol outlined in Diaz-Rojas et al. (2021)[21] for fMRI data acquisition. In summary, inside the MRI scanner, subjects were shown silent videos of a male model performing actions from the first-person perspective with a duration of 30 s (Fig 1) and were instructed to observe the movements of the model's hands. The videos were two infant-interaction videos (S1: playing with an infant; S2: changing diapers) and two control videos without infant interaction (C1: opening a box and removing a tripod from it; C2: wrapping a box with plastic). The male model and the infant were a father–child dyad of East Asian ethnicity and were unrelated to any of the participants. The control videos were performed by the same male model and were done to roughly match the movements in the corresponding infant-interaction videos. In the infant-interaction videos, the frame was set to exclude the infant's eyes, because previous studies demonstrated that the facial features of infants may evoke different responses in

human males (i.e. own- vs. other-child[58,59] due to the increased response due to facial resemblance[60,61]). Therefore, we opted to minimise the confounding effects in our subjects. The control videos were then framed accordingly. The experiment followed a block design, where each video was presented only once per run followed by 30 s of rest (static grey screen). The order of presentation was pseudo-randomised with the caveat that the infant interaction and its respective control video were always presented consecutively of each other (for example, one run would be S1–C1–S2–C2, whereas the other was S2–C2–C1–S1). Each subject completed two runs of this task. Inside the scanner, the subjects also performed an auditory task immediately after the two video runs. Data for this task were used for another experiment and were, thus, excluded from the present study.

Emotional rating of stimulus

At the end of Session 3, all participants rated each stimulus using a seven-point Likert-type scale for two attributes, namely, emotional valence and arousal. The items for emotional valence range from strong feelings of pleasantness to strong feelings of unpleasantness, whereas those for arousal range from very excited to very calm. There were no differences in the emotional valence or arousal rating between groups (Supplementary Fig S2). N-way ANOVA revealed a significant effect in the valence rating due to the type of stimulus (infant > control video, $F=140.19$, $p<0.0001$), and a smaller effect due to stimulus context (S1 or C1 vs S2 or C2, $F=5.25$, $p=0.01$). For arousal ratings, N-way ANOVA showed a significant group effect (papa vs control, $F=5.15$, $p=0.02$), and due to the type of stimulus (infant vs control video, $F=25$, $p=0.0004$).

fMRI data acquisition and processing

Neuroimaging was performed using a 3T MRI MAGNETOM Verio (Siemens Healthcare, Erlangen, Germany). Functional T2*-weighted images were collected using a gradient-Echo-Planar Imaging sequence with the following parameters: repetition time = 3000 ms, echo time = 30 ms, field of view = 192 mm, matrix size = 64 × 64, flip angle = 90°, voxel size = 3 × 3 × 3 mm; slice gap = 0; number of slices = 46 axial slices; slice order = interleaved. High-resolution structural T1-weighted images were collected using a three-dimensional magnetisation-prepared rapid acquisition with a gradient echo sequence using the following parameters: repetition time = 2250 ms, echo time = 3.51 ms, field of view = 256 mm, matrix size = 256 × 256, flip angle = 90°, voxel size = 1 × 1 × 1 mm, slice gap = 0; number of slices = 208 axial slices; slice order = interleaved.

Preprocessing of fMRI data was conducted using MATLAB R2018b (MathWorks, Natick, USA) and the SPM software package (SPM12 v7487, https://www.fil.ion.ucl.ac.uk/spm/). The first eight volumes for each run were discarded to allow for signal stabilisation. Functional images were corrected for acquisition time, resliced, realigned and normalised to match the MNI template brain and were spatially

smoothed using a Gaussian kernel with a full width at half maximum (FWHM) of 4 mm. Subjects with movement artefacts greater than 3 mm within the runs were discarded (N = 2; one PG subject for first session and one CG subject for second session).

We modelled the response of each subject to the stimuli using a general linear model in which the stimulus blocks were defined as predictors and convolved with the standardised model of hemodynamic response function, and the head motion parameters as nuisance factors. For analysis, we defined three contrasts, namely, S1–C1, S2–C2 and the overall S–C (a combination of S1 and S2 and C1 and C2). However, to avoid the unnecessary increase in the number of tests, we focused mainly on the S–C contrast and used the two other contrasts for exploratory analyses.

fMRI whole-brain analysis

Whole-brain analysis for the interactions between group and session effects were measured using a single 2 × 3 flexible factorial model (group × session), in which each participant served as a random effect. Data were drawn from the contrast map for S–C for each session and for all subjects. To verify the cross-sectional and longitudinal effects as well as intra and inter-group effects, we performed the following analysis using individual *t*- and *F*-tests.

Cross-sectional analysis

For each session (session 2 and 3), we conducted *t*-tests for the null hypothesis that the differences observed in a given voxel cluster between the PG and CG would be negligible. Two *t*-tests were conducted for each session, one for the possibility that PG > CG and one for the opposite.

Intragroup longitudinal analysis

Intragroup longitudinal comparisons (i.e. session effects) were conducted using an *F*-test for the overall session effect for each group. We also conducted *t*-tests for each pair of sessions (e.g., PG session 3 data vs session 1 data) for the null hypothesis that the differences observed in a given voxel cluster were negligible. Additionally, to test the overall development from pregnancy to postpartum, we added another *t*-test for the session 3 versus the average of session 1 & 2, for both groups.

Inter-group longitudinal analysis

Inter-group longitudinal comparisons (i.e. interaction group-session effects) were conducted using an *F*-test to determine the overall group-session effect. We also conducted *t*-tests between the two groups for each pair of sessions (e.g., PG session 3 data vs session 1 data > CG session 3 data vs session 1 data) for the null hypothesis that the differences observed in a given voxel cluster were negligible. Similar to intragroup analysis, we also tested the overall development from pregnancy to postpartum in the PG compared with the CG by adding an additional *t*-test for the session 3 versus the average of session 1 & 2 between the two groups.

Whole-brain regression analysis

To determine whether the observed neural activation in the PG and CG were related to any of the behavioural or hormonal measurements (covariates) obtained from the participants, we used whole-brain individual regression models with the S–C contrast of each group and session against each of the covariates. Diaz-Rojas et al. (2021) presented the data for session 1; thus, we did not report these data in the current paper.

Statistics for whole-brain analysis

For all whole-brain analyses, statistical maps were assessed both at (a) a family-wise error (FWE)-corrected threshold of $p < 0.05$ at the voxel level or (b) an uncorrected threshold of $p < 0.001$ at the voxel level (i.e. cluster-forming threshold), and clusters were considered significant if they passed a cluster level threshold of $p < 0.05$ after FWE correction. All significant voxels and clusters are reported. Using the WFU Pick Atlas toolbox for MATLAB[62,63] and the AAL atlas of the human brain[64], we matched the surviving clusters to known anatomical areas (Supplementary Table S1).

MVPA

MVPA was implemented using libsvm[65]. We first extracted the non-averaged and non-spatially smoothed voxel data for each subject for the S–C contrast from the left dmPFC. We focused on this ROI, because this area displayed the most consistent changes from pregnancy to postpartum. This area was anatomically defined using the left frontal superior medial area of the AAL atlas. Due to the large number of features (i.e. voxels) in this anatomical area, we performed dimensionality reduction to improve the performance of the analysis[66]. Dimensionality reduction was conducted using the one-way ANOVA *F*-values ranking of each voxel, with the target labels of the training set (group or session) as the levels, selecting the top 125 voxels (equivalent to a 5x5x5 voxels cube) with the highest *F*-values. Using this data, we trained a support-vector machine (SVM) classifier with a linear kernel to perform a binary classification of a target label. These labels were of two kinds: session (early-pregnancy, late-pregnancy

or postpartum), and paternity status (same as group, either PG or CG). To validate the model, we performed leave-one-out cross-validation. In other words, we trained a model using voxel data from all subjects included in the analysis except for one and used the excluded subject as test data. This procedure was repeated for each subject until all subjects served once as test data. Then, we averaged the prediction accuracy for each of the test data to obtain the mean prediction accuracy of the classifier for that analysis. We also displayed the 95% confidence interval, which was calculated using the normal approximation method for the binomial confidence interval. High accuracy implied that the model could separate multivoxel data between subjects belonging to one group/session from the other. In this manner, we established a relationship between neural activation patterns and the target label. Significance was assessed with an upper-tailed binomial test for the null hypothesis that the resultant classification accuracy emerged from a binomial distribution with a mean equal to the chance level (50%). Additionally, we calculated the true positive rate (TPR) and true negative rate (TNR) for each ROI–target label pair.

Exploratory analysis of the individual differences in the paternal brain

Clustering analysis was conducted using a *k*-means algorithm and the squared Euclidean distance as the distance metric. We set the number of groups as two (k=2). We intended to group the subjects based on the development of the paternal brain across pregnancy. Thus, we selected the percentage change between adjacent sessions as data for clustering (i.e. change from session 1 to 2, and from session 2 to 3). Next, we took the absolute values of these percentage change. Negative or positive changes cannot be easily interpreted as improvements or regressions in the paternal brain (for example, a negative change could be associated with decreased response due to familiarity with infant-interaction schemes or to perceived non-familiarity to the shown infant, which triggers own-child/other-child effects). Thus, we focused solely on change as a metric for the development of the parental brain. However, measuring the percentage change may frequently lead to extremely large values, because fMRI data may have very small values. To account for this phenomenon, we took the base 10 logarithm of the absolute percentage change, and then clipped to the outlier threshold any outlier values (defined as those values that are more than three scaled median absolute deviations away from the median). We focused on the neural data from the left dmPFC, which was the brain region that exhibited greater changes from pregnancy to postpartum. This analysis was conducted only with data from the PG.

The *k*-means algorithm is sensitive to initial conditions and to the initial value of the random number generator. To combat this issue and validate the findings, we first performed the following procedure. We excluded four random subjects from the set (12.5% of the total set) and performed *k*-means classification with the remaining data. We noted the cluster centres and repeated this procedure 10,000 times. Finally, we averaged the cluster centres over the iterations and performed one final clustering using these centres and the entire data set. For the final clustering, each data point was set as belonging to the group that it is closer to using the square Euclidean distance and without updating the position of the cluster centres, because this process is iteratively conducted in the *k*-means algorithm.

After dividing the subjects into two groups, we calculated for the mean of each behavioural and hormonal covariate at each time point. Statistical analysis at the group and session levels was performed using a N-Way ANOVA. Supplementary Table S3 presents the results for all tests. Behavioural data that remained stable across sessions (i.e. age, years of education, weekly worktime and household income) were taken from session 3.

We repeated the above procedures with k=3 (Supplementary Fig. S3) and k=4 (Supplementary Fig. S4) groups to confirm that our results were not impacted by the arbitrary selection of the k parameter. In the case of k=3, the results were consistent with those found with k=2. One group was exactly the same as before (Group 2 in Fig 4; Group 2 in Fig S4); the remaining two groups showed no differences in their behavioural or hormonal traits, suggesting that the division is inconsequential and no better than the simplest division of k=2. In the case of k=4, one group had only 4 subjects in it, making the interpretation of any results difficult due to lack of samples. Using k>4 results in similar small-sized groups.

References

1. Leventhal, B. L. The Role of the Father in Child Development. *J. Am. Acad. Child Psychiatry* **22**, 582–583 (1983).

2. Williams, E. & Radin, N. Effects of father participation in child rearing: Twenty-year follow-up. *Am. J. Orthopsychiatry* **69**, 328–336 (1999).

3. Feldman, R. Parents' convergence on sharing and marital satisfaction, father involvement, and parent-child relationship at the transition to parenthood. *Infant Ment. Health J.* **21**, 176–191 (2000).

4. Lamb, M. *The role of the father in child development, 5th ed. The role of the father in child development, 5th ed.* (John Wiley & Sons Inc, 2010).

5. Lamb, M. E. *The Father's role : cross-cultural perspectives*. (Hillsdale, NJ : L. Erlbaum Associates, 1987).

6. Kramer, K. L. Cooperative Breeding and its Significance to the Demographic Success of Humans. *Annu. Rev. Anthropol.* **39**, 417–436 (2010).

7. Shackelford T & Salmon C. *The Oxford Handbook of Evolutionary Family Psychology*. (Oxford University Press, 2011). doi:10.1093/oxfordhb/9780195396690.001.0001.

8. Feldman, R., Braun, K. & Champagne, F. A. The neural mechanisms and consequences of paternal caregiving. *Nat. Rev. Neurosci.* **20**, 205–224 (2019).

9. Sarkadi, A., Kristiansson, R., Oberklaid, F. & Bremberg, S. Fathers' involvement and children's developmental outcomes: a systematic review of longitudinal studies. *Acta Paediatr.* **97**, 153–158 (2008).

10. Kentner, A. C., Abizaid, A. & Bielajew, C. Modeling Dad: Animal models of paternal behavior. *Neurosci. Biobehav. Rev.* **34**, 438–451 (2010).

11. Swain, J. E., Dayton, C. J., Kim, P., Tolman, R. M. & Volling, B. L. Progress on the paternal brain: theory, animal models, human brain research, and mental health implications. *Infant Ment. Health J.* **35**, 394–408 (2014).

12. Abraham, E. *et al.* Father's brain is sensitive to childcare experiences. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9792–9797 (2014).

13. Provenzi, L. *et al.* The paternal brain in action: A review of human fathers' fmri brain responses to child-related stimuli. *Brain Sci.* **11**, (2021).

14. Isik, L., Koldewyn, K., Beeler, D. & Kanwisher, N. Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9145–E9152 (2017).

15. Rilling, J. K. & Mascaro, J. S. The neurobiology of fatherhood. *Curr. Opin. Psychol.* **15**, 26–32 (2017).

16. Craig, A. D. B. How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* **10**, (2009).

17. Fan, Y., Duncan, N. W., de Greck, M. & Northoff, G. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neurosci. Biobehav. Rev.* **35**, 903–911 (2011).

18. Rizzolatti, G. & Craighero, L. The Mirror-Neuron System. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).

19. Lenzi, D. *et al.* Neural basis of maternal communication and emotional expression processing during infant preverbal stage. *Cereb. Cortex* **19**, 1124–1133 (2009).

20. Seifritz, E. *et al.* Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. *Biol. Psychiatry* **54**, 1367–1375 (2003).

21. Diaz-Rojas, F. *et al.* Development of the paternal brain in expectant fathers during early pregnancy. *Neuroimage* **225**, (2021).

22. Kim, P. *et al.* Neural plasticity in fathers of human infants. *Soc. Neurosci.* **9**, 522–535 (2014).

23. Berg, S. J. & Wynne-Edwards, K. E. Changes in Testosterone, Cortisol, and Estradiol Levels in Men Becoming Fathers. *Mayo Clin. Proc.* **76**, 582–592 (2001).

24. Edelstein, R. S. *et al.* Prospective and dyadic associations between expectant parents' prenatal hormone changes and postpartum parenting outcomes. *Dev. Psychobiol.* **59**, 77–90 (2017).

25. Saxbe, D. E., Schetter, C. D., Simon, C. D., Adam, E. K. & Shalowitz, M. U. High paternal testosterone may protect against postpartum depressive symptoms in fathers, but confer risk to mothers and children. *Horm. Behav.* **95**, 103–112 (2017).

26. Gordon, I., Zagoory-Sharon, O., Leckman, J. F. & Feldman, R. Oxytocin and the Development of Parenting in Humans. *Biol. Psychiatry* **68**, 377–382 (2010).

27. Weisman, O., Zagoory-Sharon, O. & Feldman, R. Oxytocin administration, salivary testosterone, and father–infant social behavior. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **49**, 47–52 (2014).

28. Kuo, P. X. *et al.* Fathers' cortisol and testosterone in the days around infants' births predict later paternal involvement. *Horm. Behav.* **106**, 28–34 (2018).

29. Feldman, R., Gordon, I., Schneiderman, I., Weisman, O. & Zagoory-Sharon, O. Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent-infant contact. *Psychoneuroendocrinology* **35**, 1133–1141 (2010).

30. Leibenluft, E., Gobbini, M. I., Harrison, T. & Haxby, J. V. Mothers ' Neural Activation in Response to Pictures of Their Children and Other Children. 225–232 (2004) doi:10.1016/j.biopsych.2004.05.017.

31. Nitschke, J. B. *et al.* Orbitofrontal cortex tracks positive mood in mothers viewing pictures of their newborn infants. **21**, 583–592 (2004).

32. Wittfoth-Schardt, D. *et al.* Oxytocin modulates neural reactivity to children's faces as a function of social salience. *Neuropsychopharmacology* **37**, 1799–1807 (2012).

33. Young, K. S. *et al.* The neural basis of responsive caregiving behaviour: Investigating temporal dynamics within the parental brain. *Behav. Brain Res.* **325**, 105–116 (2017).

34. Swain, J. E. The human parental brain: In vivo neuroimaging. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **35**, 1242–1254 (2011).

35. Rilling, J. K. The neural and hormonal bases of human parentalcare. *Neuropsychologia* **51**, 731–747 (2013).

36. Cataudella, S., Lampis, J., Busonera, A., Marino, L. & Zavattini, G. C. From parental-fetal attachment to a parent-infant relationship : A systematic review about prenatal protective and risk factors From parental-fetal attachment to a parent-infant relationship : a systematic review about prenatal protective and risk factors. (2016).

37. Inori, S. & Kato, M. Attitude toward childrearing and the relevant factors of female adolescents: influences of views on sex roles, feelings to children, and attitude toward childrearing of mother (original in Japanese). *Proc. Fac. Hum. Cult. Hiroshima Univ.* **6**, 199–208 (2011).

38. Vreeswijk, C. M. J. M., Maas, A. J. B. M., Rijk, C. H. A. M., Braeken, J. & van Bakel, H. J. A. Stability of fathers' representations of their infants during the transition to parenthood. *Attach. Hum. Dev.* **16**, 292–306 (2014).

39. Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc. Natl. Acad. Sci.* **98**, 4259–4264 (2001).

40. Piva, M. *et al.* The dorsomedial prefrontal cortex computes task-invariant relative subjective value for self and other. *Elife* **8**, 1–29 (2019).

41. Ford, J. H. & Kensinger, E. A. Older adults recruit dorsomedial prefrontal cortex to decrease negativity during retrieval of emotionally complex real-world events. *Neuropsychologia* **135**, 107239 (2019).

42. Johnson, S., Li, J., Kendall, G., Strazdins, L. & Jacoby, P. Mothers' and Fathers' Work Hours, Child Gender, and Behavior in Middle Childhood. *J. Marriage Fam.* **75**, 56–74 (2013).

43. Bünning, M. Paternal Part-Time Employment and Fathers' Long-Term Involvement in Child Care and Housework. *J. Marriage Fam.* **82**, 566–586 (2020).

44. Guillot, A. *et al.* Brain activity during visual versus kinesthetic imagery: An fMRI study. *Hum. Brain Mapp.* **30**, 2157–2172 (2009).

45. Lindner, I., Schain, C. & Echterhoff, G. Other-self confusions in action memory: The role of motor processes. *Cognition* **149**, 67–76 (2016).

46. Gettler, L. T., McDade, T. W., Feranil, A. B. & Kuzawa, C. W. Longitudinal evidence that fatherhood decreases testosterone in human males. *Proc. Natl. Acad. Sci.* **108**, 16194–16199 (2011).

47. Spielberger, C. D. State-Trait Anxiety Inventory. in *The Corsini Encyclopedia of Psychology* (John Wiley & Sons, Inc., 2010). doi:10.1002/9780470479216.corpsy0943.

48. Zich, J. M., Attkisson, C. C. & Greenfield, T. K. Screening for depression in primary care clinics: The CES-D and the BDI. *Int. J. Psychiatry Med.* **20**, 259–277 (1990).

49. Wee, K. Y., Skouteris, H., Pier, C., Richardson, B. & Milgrom, J. Correlates of ante- and postnatal depression in fathers: A systematic review. *J. Affect. Disord.* **130**, 358–377 (2011).

50. Wilson, S. & Durbin, C. E. Effects of paternal depression on fathers' parenting behaviors: A meta-analytic review. *Clin. Psychol. Rev.* **30**, 167–180 (2010).

51. Spanier, G. B. Assessing the strengths of the Dyadic Adjustment Scale. *J. Fam. Psychol.* **2**, 92–94 (1988).

52. Condon, J. T. The assessment of antenatal emotional attachment: Development of a questionnaire instrument. *Br. J. Med. Psychol.* **66**, 167–183 (1993).

53. Morishita, Y. The Effect of Becoming Fathers on Men's Development. *Japanese J. Dev. Psychol.* **17**, 182 (2006).

54. Condon, J. T., Corkindale, C. J. & Boyce, P. Assessment of postnatal paternal–infant attachment: development of a questionnaire instrument. *J. Reprod. Infant Psychol.* **26**, 195–210 (2008).

55. Volling, B. L. & Belsky, J. The Contribution of Mother-Child and Father-Child Relationships to the Quality of Sibling Interaction: A Longitudinal Study. *Child Dev.* **63**, 1209 (1992).

56. Shimozuru, M., Kikusui, T., Takeuchi, Y. & Mori, Y. Effects of isolation-rearing on the development of social behaviors in male Mongolian gerbils (Meriones unguiculatus). *Physiol. Behav.* **94**, 491–500 (2008).

57. Aoki, M., Shimozuru, M., Kikusui, T., Takeuchi, Y. & Mori, Y. Sex Differences in Behavioral and Corticosterone Responses to Mild Stressors in ICR Mice are Altered by Ovariectomy in Peripubertal Period. *Zoolog. Sci.* **27**, 783–789 (2010).

58. Kuo, P. X., Carp, J., Light, K. C. & Grewen, K. M. Neural responses to infants linked with behavioral interactions and testosterone in fathers. *Biol. Psychol.* **91**, 302–306 (2012).

59. Mascaro, J. S., Hackett, P. D. & Rilling, J. K. Testicular volume is inversely correlated with nurturing-related brain activity in human fathers. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15746–15751 (2013).

60. Platek, S. M. *et al.* Reactions to children's faces: Males are more affected by resemblance than females are, and so are their brains. *Evol. Hum. Behav.* **25**, 394–405 (2004).

61. Platek, S. M., Keenan, J. P. & Mohamed, F. B. Sex differences in the neural correlates of child facial resemblance: An event-related fMRI study. *Neuroimage* **25**, 1336–1344 (2005).

62. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19**, 1233–1239 (2003).

63. Maldjian, J. A., Laurienti, P. J. & Burdette, J. H. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage* **21**, 450–455 (2004).

64. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).

65. Chang, C.-C. & Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 1–27 (2011).

66. Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D. & Brovelli, A. Multivoxel Pattern Analysis for fMRI Data: A Review. *Comput. Math. Methods Med.* **2012**, 1–14 (2012).

Tables

Table 1. Summary of behavioural and hormonal traits. ANOVA *p* values shown are with correction for number of comparisons. Values shown are mean ± standard deviation. *No intergroup statistical test was conducted for *Corrected DAS*, due to lack of married subjects in CG.

	Father			Control			Anova p values		
Covariate	1st	2nd	3rd	1st	2nd	3rd	Group	Session	Group x Session
N	33	35	34	35	33	33	-	-	-
Age (years)	33.76 ± 5.31	33.67 ± 5.51	34.15 ± 5.57	31.26 ± 6.13	32.07 ± 6.01	32.34 ± 6.08	0.24	1.00	1.00
BDI	5.79 ± 4.59	4.69 ± 4.13	6.06 ± 6.10	8.17 ± 6.92	6.55 ± 5.79	6.55 ± 5.12	0.65	1.00	1.00
Corrected DAS*	0.75 ± 0.09	0.77 ± 0.07	0.75 ± 0.09	0.69 ± 0.08	0.62 ± 0.19	0.71 ± 0.06	-	-	-
Education (years)	17.00 ± 2.05	17.00 ± 2.00	17.03 ± 2.02	16.89 ± 3.06	16.42 ± 4.29	16.91 ± 3.13	1.00	1.00	1.00
Foetal/Infant Attachment	0.74 ± 0.14	0.79 ± 0.11	0.80 ± 0.09	- ± -	- ± -	- ± -	-	-	-
GA (weeks)	20.18 ± 5.50	34.74 ± 1.84	56.26 ± 3.37	- ± -	- ± -	- ± -	-	-	-
Degree of house chores	0.58 ± 0.16	0.59 ± 0.12	0.56 ± 0.14	0.59 ± 0.07	0.56 ± 0.09	0.62 ± 0.06	1.00	1.00	1.00
Household income	6.61 ± 2.46	6.43 ± 2.52	6.50 ± 2.53	2.66 ± 2.13	2.82 ± 2.08	2.76 ± 2.14	8.97E-23	1.00	1.00
Negative attitude towards parenting	3.85 ± 0.56	3.80 ± 0.50	3.66 ± 0.59	3.90 ± 0.47	3.77 ± 0.52	3.84 ± 0.64	1.00	1.00	1.00
Oxytocin (pg/ml)	403.81 ± 321.5	244.61 ± 159.3	238.24 ± 216.2	348.41 ± 295.7	212.95 ± 188.2	201.52 ± 209.6	1.00	2.37E-03	1.00
Oxytocin (residuals)	0.48 ± 1.33	-0.09 ± 0.64	-0.19 ± 0.85	0.26 ± 1.18	-0.22 ± 0.77	-0.24 ± 0.86	1.00	0.01	1.00
Positive attitude towards parenting	4.44 ± 0.47	4.52 ± 0.38	4.64 ± 0.43	4.18 ± 0.49	4.17 ± 0.62	4.14 ± 0.73	2.21E-05	1.00	1.00
State anxiety score	36.39 ± 7.05	36.26 ± 7.49	34.68 ± 7.63	35.63 ± 5.82	35.73 ± 6.28	37.73 ± 7.78	1.00	1.00	1.00
Testosterone (pg/ml)	0.45 ± 0.24	0.48 ± 0.22	0.51 ± 0.27	0.37 ± 0.19	0.52 ± 0.24	0.56 ± 0.26	1.00	0.09	1.00
Testosterone (residuals)	-0.03 ± 1.13	0.11 ± 0.90	-0.04 ± 1.07	-0.39 ± 0.88	0.18 ± 1.00	0.19 ± 0.90	1.00	1.00	1.00
Trait anxiety score	38.79 ± 9.57	39.80 ± 10.18	37.50 ± 10.65	43.71 ± 12.10	43.21 ± 10.14	44.24 ± 9.54	0.01	1.00	1.00
Weekly Worktime (hours)	50.42 ± 12.16	50.54 ± 11.95	50.56 ± 12.13	28.49 ± 16.32	30.09 ± 15.38	28.88 ± 16.13	2.40E-20	1.00	1.00
Fathers development	-	-	4.16 ± 0.43	-	-	-	-	-	-
Postnatal attachment	-	-	0.80 ± 0.09	-	-	-	-	-	-
Parental involvement	-	-	0.61 ± 0.09	-	-	-	-	-	-
Time alone with child (hours/week)	-	-	6.74 ± 8.53	-	-	-	-	-	-
Infant age (weeks)	-	-	17.64 ± 4.25	-	-	-	-	-	-
GA at birth	-	-	38.66 ± 3.24	-	-	-	-	-	-

Table 2. Significant clusters of activations for stimulus versus control condition for cross-sectional comparisons. Only clusters with a family-wise error rate of $p<0.05$ (pFWE) at the cluster or peak level are shown. False discovery rate-corrected (FDR) p values are also shown. Areas were defined based on AAL neuroatlas. XYZ coordinates are in MNI space. p unc.: Uncorrected p-values. T: t-test; PG: father group; CG: control group. +: Positive effect (the higher the brain activity, the higher the covariate value); -: negative effect (the higher the brain activity, the lower the covariate value).

Cross-sectional comparisons: Experimental condition versus control condition (S > C)

Session/contrast	*p* FWE (peak)	*p* FWE (cluster)	*p* FDR (cluster)	*p* unc. (cluster)	Cluster size	T / *F* Statistic	X	Y	Z	ROI
T – Session 2: CG > PG	0.031	0.022	0.033	0.002	28	4.92	6	62	17	Frontal_Sup_Medial_R
T – Session 3: PG > CG	0.034	0.003	0.008	<0.001	42	4.89	−6	56	5	Frontal_Sup_Medial_L

Cross-sectional comparisons: S > C contrast versus covariates

Covariates	Group	Session	Effect	*p* FWE (peak)	*p* FDR (cluster)	*p* unc. (cluster)	Cluster size	X	Y	Z	ROI
Oxytocin (no nuis)	PG	2	−	0.023	0.124	0.008	15	12	−4	68	Supp_Motor_Area_R
Household income	CG	2	+	0.027	0.449	0.096	5	−54	−1	38	Precentral_L
Age	PG	3	+	0.179	0.002	<0.001	57	39	38	20	Frontal_Mid_R
DAS	PG	3	−	0.029	0.423	0.021	12	6	5	65	Supp_Motor_Area_R
DAS	PG	3	−	0.049	0.423	0.027	11	15	14	62	Supp_Motor_Area_R
DAS	PG	3	−	0.617	0.05	<0.001	35	3	−16	32	Cingulum_Mid_R
House chores	PG	3	−	0.006	0.488	0.029	11	63	−46	−7	Temporal_Mid_R
Postnatal Attach	PG	3	−	0.011	<0.001	<0.001	102	24	35	35	Frontal_Sup_R
Postnatal Attach	PG	3	−	0.085	0.027	0.001	26	6	−16	35	Cingulum_Mid_R
Postnatal Attach	PG	3	−	0.205	<0.001	<0.001	70	−3	−64	53	Precuneus_L
Postnatal Attach	PG	3	−	0.296	0.002	<0.001	46	−21	41	32	Frontal_Mid_L
Postnatal Attach	PG	3	−	0.463	0.002	<0.001	49	54	−40	41	SupraMarginal_R
Postnatal Attach	PG	3	−	0.508	0.001	<0.001	52	12	−25	44	Cingulum_Mid_R
Postnatal Attach	PG	3	−	0.561	0.005	<0.001	38	9	−49	47	Precuneus_R
Postnatal Attach	PG	3	−	0.681	0.035	0.002	23	−33	14	2	Insula_L
Postnatal Attach	PG	3	−	0.902	0.035	0.002	23	6	44	11	Cingulum_Ant_R

Table 3. Significant clusters of activations for the stimulus versus control condition for all longitudinal comparisons. Only clusters with a FWE rate of $p < 0.05$ (pFWE) at the cluster or peak level are shown. False discovery rate-corrected (FDR) p values are also shown. Areas were defined based on the AAL neuroatlas. The XYZ coordinates are in MNI space. p unc.: Uncorrected p values. T: *t*-test; F: *F*-test; PG: father group; CG: control group.

Longitudinal comparisons:										
Session / contrast	*p*FWE (peak)	*p*FWE (cluster)	*p*FDR (cluster)	*p* unc. (cluster)	Cluster size	T/*F* Statistic	X	Y	Z	ROI
F - Group - Session Interaction	0.009	0.002	0.001	0	35	16.66	−3	59	8	Frontal_Sup_Medial_L
F - Group - Session Interaction	0.472	0.081	0.036	0.005	15	11.73	6	62	20	Frontal_Sup_Medial_R
T - Papa Session 3 > Session 1 & 2	0.092	0	0	0	78	4.66	−3	53	11	Frontal_Sup_Medial_L
T - Papa Session 3 > Session 2	0.027	0.729	0.322	0.104	6	4.94	42	8	−28	Temporal_Pole_Sup_R
T - Papa Session 3 > Session 2	0.156	0	0	0	67	4.53	−6	44	26	Frontal_Sup_Medial_L
T – Session 3 > Session 1 - Papa > Control	0.985	0.066	0.044	0.005	21	3.57	−3	47	17	Frontal_Sup_Medial_L
T – Session 3 > Session 2 - Papa > Control	0.001	0	0	0	74	5.77	−3	59	8	Frontal_Sup_Medial_L
T – Session 3 > Session 2 - Papa > Control	0.126	0.005	0.006	0	38	4.58	6	62	17	Frontal_Sup_Medial_R
T – Session 3 > Session 1&2 - Papa > Control	0.001	0	0	0	72	5.07	−3	59	8	Frontal_Sup_Medial_L

Figures

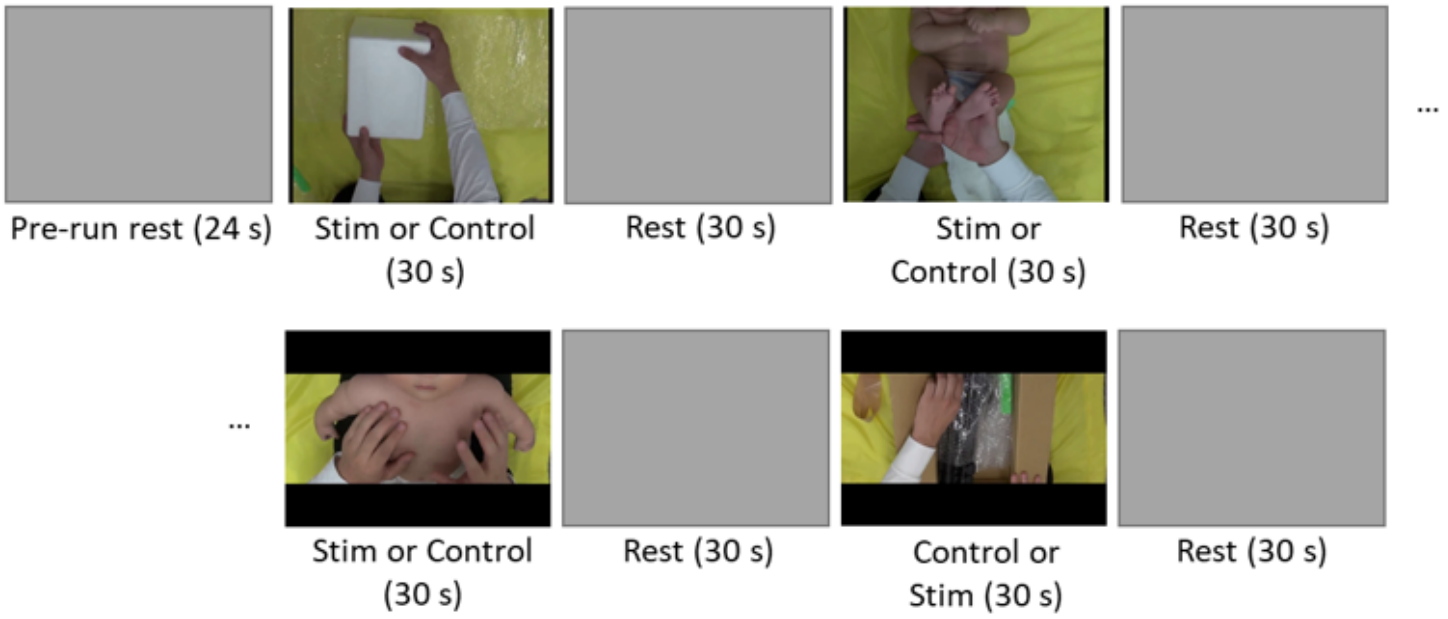

Figure 1

fMRI experimental design: an example of a typical run inside the MRI scanner. The participants were instructed to view the alternating videos for infant interaction or matched control. Videos were played for 30 s with 30 s of rest. The order of presentation was pseudo-random.

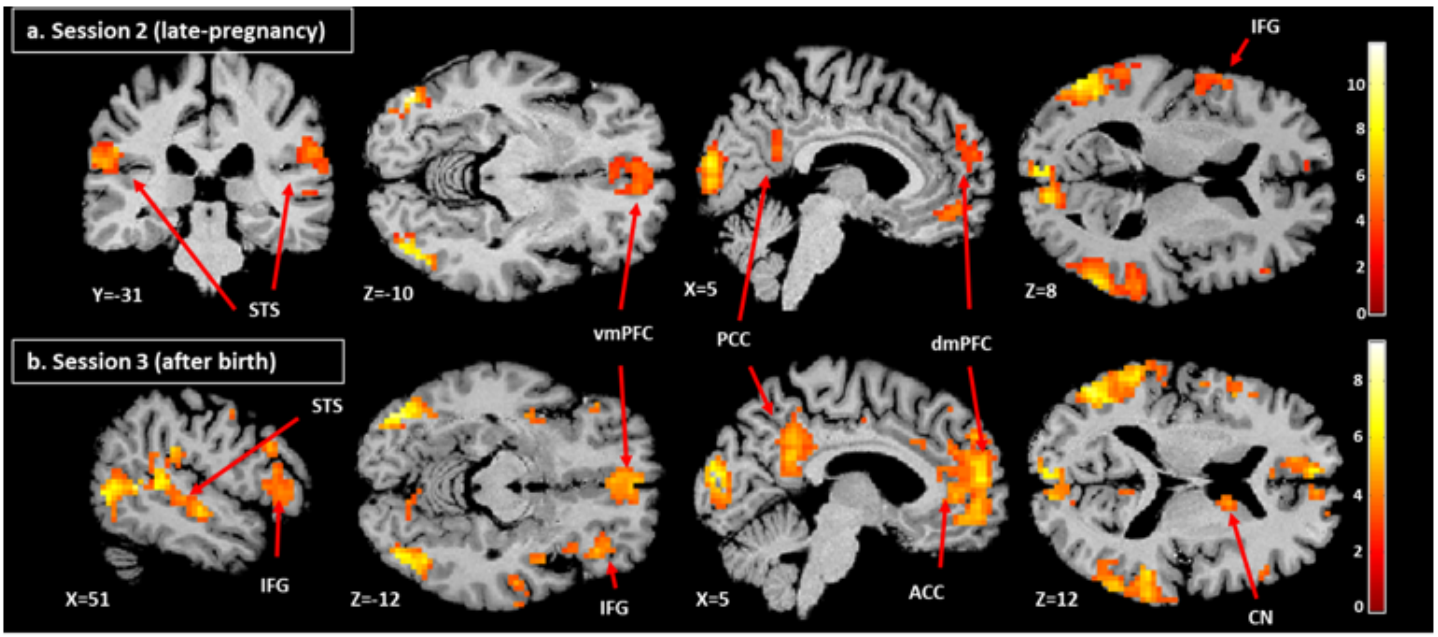

Figure 2

Whole-brain analysis showing the T contrast for S > C across conditions. Activation in the S > C contrast based on all participants for the recording session 2 (top) and 3 (bottom). The activation maps were thresholded at $p < 0.001$ (uncorrected at the voxel level) combined with $p < 0.05$ (FWE-corrected at the cluster level). ACC: anterior cingulate cortex; CN: caudate nucleus; IFG: inferior frontal gyrus; PCC: posterior cingulate cortex; STS: superior temporal sulcus; dmPFC: dorsomedial prefrontal cortex and vmPFC: ventromedial prefrontal cortex.

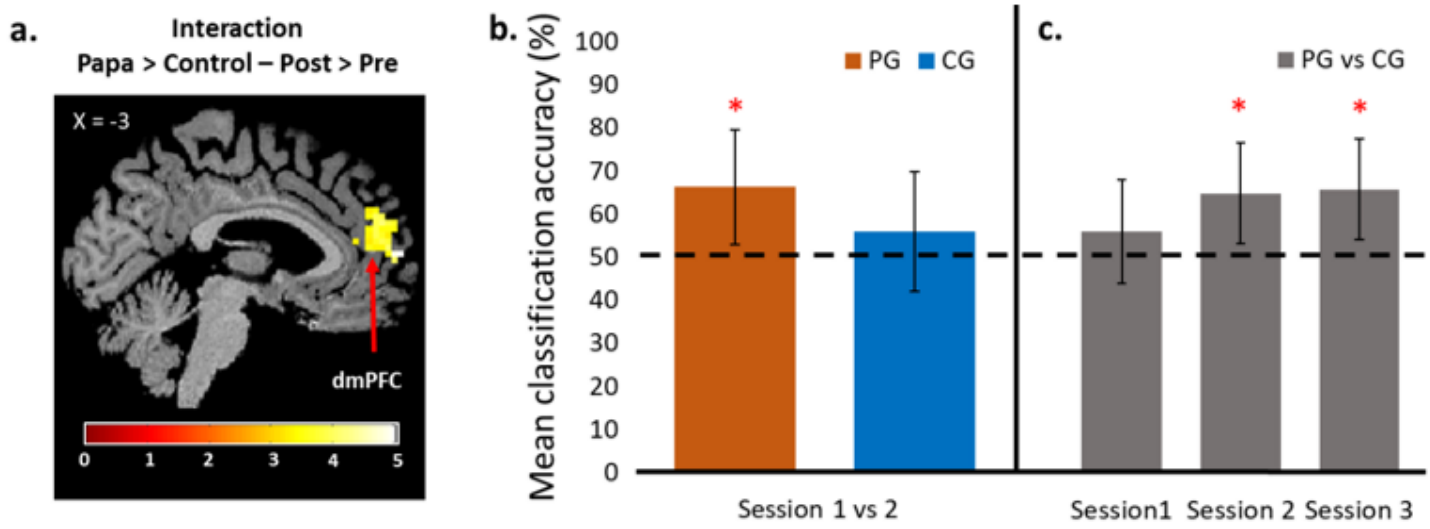

Figure 3

Changes in the paternal brain across pregnancy and after childbirth. A. Interaction between time (session 3 vs sessions 1 and 2) and group, which displays the clusters in which fathers showed an increase in

activity from sessions 1 and 2 to session 3 compared with control. The activation maps were thresholded at $p < 0.001$ (uncorrected at the voxel level) combined with $p < 0.05$ (FWE-corrected at the cluster level). dmPFC = dorsomedial prefrontal cortex. B. Classification of recording session and paternal status using multivoxel data. Using a SVM classifier, we successfully discriminated between fathers in the early pregnancy and those in late-pregnancy from the multivoxel neural patterns in the left dorsomedial prefrontal cortex. C. Using the same procedure as in B, we discriminated between fathers (PG) and control (CG) based on neural patterns from the second and third recording sessions. Error bars denote 95% CI. *$p < 0.05$.

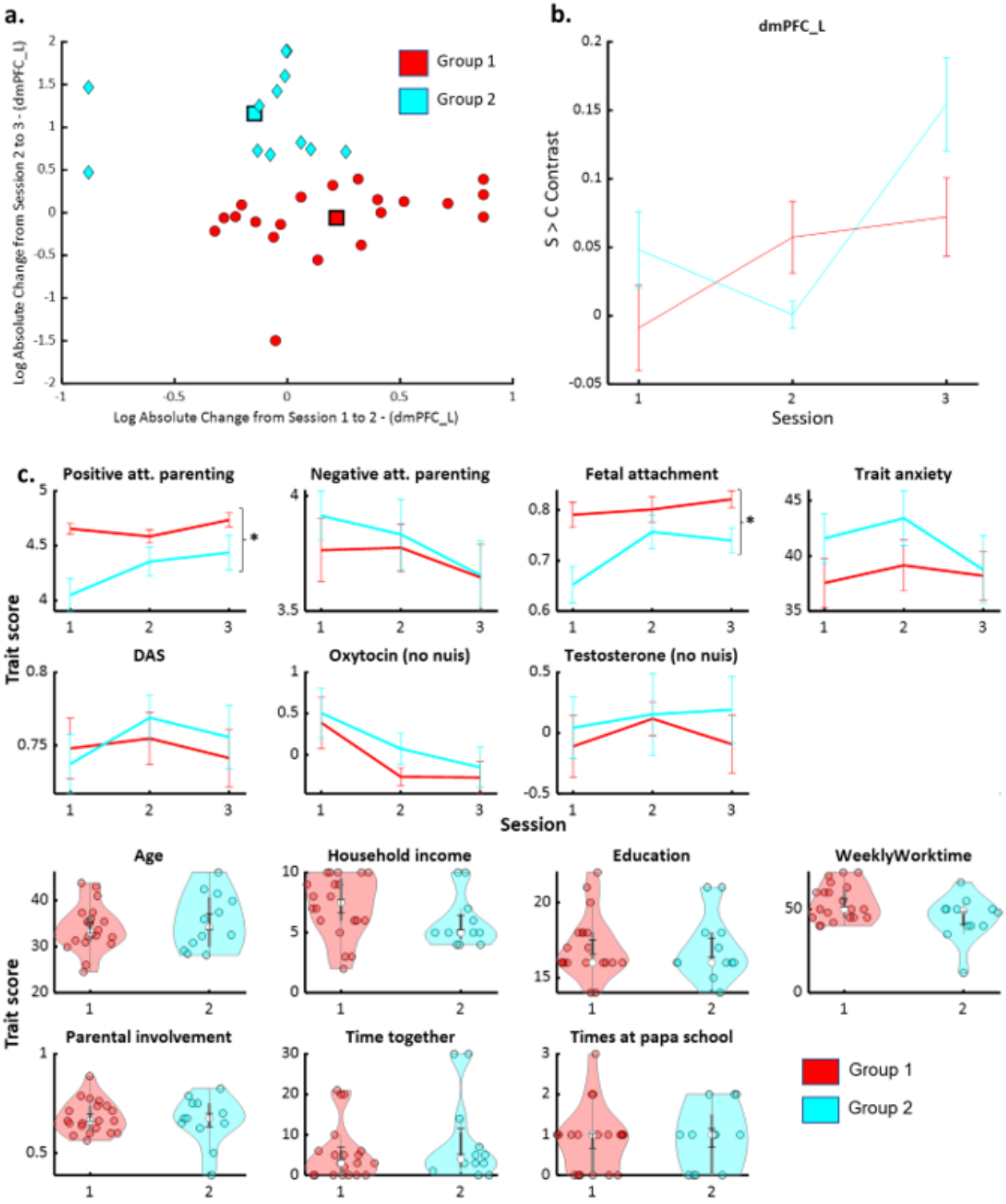

Figure 4

Clustering analysis of change in the left dorsomedial prefrontal cortex (dmPFC) in expectant fathers between sessions. A. Scatter plot of the change in dmPFC activation between sessions 1 and 2 versus between sessions 2 and 3. Change was measured using the absolute logarithm of the percentage change between the two sessions. Squares in bold represent the cluster average centres, which are calculated from 10,000 iterations of clustering using 87.5% of the total sample of fathers. B. Line plots showing the

S > C contrast signal for the dmPFC for both groups. C. Line and violin plots for each covariate divided by groups. *Group differences ($p < 0.001$, uncorrected). Error bars in line plots correspond to standard error of the mean. In the violin plots, each coloured dot corresponds to one subject's data; white dots represent the median of each group; the thick dark lines inside each violin plot represent the range between the first and third quartile.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary211129.pptx
- Supplementarydiscussion.docx