Motherhood and affect recognition: mothers need more information to correctly recognise emotions of children

Plank, Irene Sophia^{1,2,3,4 *}; Christiansen, Lina-Nel⁵; Kunas, Stefanie L.¹;

Dziobek, Isabel^{2,3,4}; Bermpohl, Felix^{1,2,3}

Running title: Motherhood and affect recognition

¹Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

²Einstein Center for Neurosciences, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

³Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

⁴Department of Psychology, Institute of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany

⁵Center for Chronically Sick Children, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

* Corresponding author

E-mail: irene-sophia.plank@charite.de

18 **Abstract**

19 Infant and child faces receive more attention than adult faces, especially when they are portraying
20 emotions. This preferential treatment is increased in mothers. It is unclear how this affects the
21 ability to recognise child emotions and at which age children lose this preferential treatment. Using
22 morphed videos where a neutral face continually and gradually takes on an emotion, we were able
23 to quantify the amount of information needed to identify the emotion. We investigated amount of
24 information needed as well as accuracy both separately and as an integrated performance score. We
25 expected mothers to perform better than non-mothers and both groups to perform better when
26 encountering adolescents as compared to children. We also hypothesised an interaction with the
27 difference between mothers and non-mothers being larger for pre-schoolers than for adolescents.
28 Both groups indeed performed better when recognising the emotions of adolescents than of pre-
29 schoolers possibly due to their increasingly adult features. Contrary to our hypotheses, mothers
30 performed worse than non-mothers in every age group due to their need for more information to
31 identify the emotions with the same level as accuracy as non-mothers. Both groups performed
32 similarly on a control task where they had to identify animals instead of emotions. This indicates that
33 mothers need more information for affect recognition specifically and not for similar tasks not
34 involving emotions. A possible explanation is that emotional child faces have a greater emotional
35 impact on mothers than non-mothers leading to task interference. Further research is needed to
36 examine possible causes of this effect including differences in affective social understanding as well
37 as whether the effect extends to adult faces.

38 Introduction

39 Parenting presents humans with a challenge: they are responsible for understanding and caring
40 for another human being that starts off as completely dependent on them [1]. This requires them to
41 be able to read social cues and understand their offspring even when their children are not able to
42 properly express themselves, either because they are preverbal or because they are overwhelmed
43 with this task [2]. Therefore, social understanding is vital for parenting [3–5].

44 Many aspects of social understanding rely on people correctly recognising the emotions of their
45 counterpart. Nonetheless, research on how motherhood influences affect recognition is scarce.
46 Matsunaga and colleagues used morphed pictures of infants and adults to determine mothers' and
47 nulliparous women's sensitivity in an affect recognition task [6]. They found no effect of motherhood
48 on sensitivity; however, since only primiparous mothers of infants less than one year old took part in
49 the study, Matsunaga and colleagues raised the possibility that differences between mothers and
50 non-mothers in affect recognition might still develop. Nishitani and colleagues used functional near-
51 infrared spectroscopy (fNIRS) to investigate neural activity during affect recognition in mothers of
52 children between 9 and 36 months old and matched non-mothers using adult and infant faces as
53 stimuli [7]. They found no behavioural differences between mothers and non-mothers, but they
54 found increased activity in response to infant stimuli in the mothers in the prefrontal cortex.
55 Similarly, Plank and colleagues found increased activity in emotion processing and theory of mind
56 areas in mothers compared to non-mothers when predicting action intentions based on facial affect
57 in children and adults [8]. They reported no behavioural differences but noted that their task was
58 designed to increase detection in neural differences and not for sensitivity on the behavioural level.
59 This increase in neural activation found in several studies could be a precursor of behavioural
60 changes in affect recognition in mothers that could come to light in a sample of women who have
61 been mothers for a longer period of time using sensitive behavioural measures.

Infant and child faces are treated differently than adult faces. Infant's emotional signals can be ambiguous and therefore harder to understand than other adults' [9] and research has revealed an attentional bias toward infant and child faces in both mothers and non-mothers which may serve to counteract this difficulty [10]. However, it is unclear at which point during childhood this preferential treatment which is often referred to as Kindchenschema is lost [11]. Attractiveness and likeability of infant and child faces seem to linearly decrease from infants to 7 year old children while typicality increases [12]. Thompson-Booth and colleagues showed that affective infant faces (6-12 months old) interfered with an unrelated task, however, this was not the case for adolescent (13-16 years old) and adult faces [13]. Child faces between 4 and 7 years of age presented a middle ground with less interference than infant faces but more than adolescent and adult faces. However, there were no stimuli of children between 7 and 13 years old. Taken together, these studies suggest that young children have a greater emotional effect on adults and are harder to read but this difference is lost at some point during childhood or adolescence when features become increasingly adult. It is still unclear when this shift happens in relation to affect recognition.

The goal of the present study is to extend our understanding of how motherhood influences affect recognition of unfamiliar children from pre-schoolers to adolescents. We focussed on motherhood because in Germany mothers are more likely to assume primary care of their children than fathers [14]. We used stimuli of children between 4 and 15 years old in three age groups (pre-schoolers, middle children and adolescents). Mothers always encountered stimuli in the age group closest to their child's or children's age. Since affect recognition is a fundamental task, it is vital to use a paradigm that is sensitive enough to pick up subtle differences. We created morphed videos that show a gradual and continuous change from a neutral face to a face expressing an emotion (afraid, angry, happy, sad). Participants were asked to stop the videos as soon as they recognised the emotion, thereby giving us an estimate of how much information they needed to correctly identify the emotion. This response time was combined with their accuracy to a linear integrated speed accuracy score (LISAS) to measure their performance [15]. We used a similar task where neutral

faces morphed into animals as a control. We hypothesised that mothers would perform better in facial affect recognition of an unfamiliar child than non-mothers, but that mothers and non-mothers would not differ in the control task. Additionally, we expected higher performance in the oldest age group due to adolescent features becoming increasingly similar to adult features. We also hypothesised that motherhood would interact with the age of the stimuli with larger differences between mothers and non-mothers in pre-schoolers. Additionally, we explored the effect of emotion type on both age group and motherhood.

Methods

Participants

Participants for this study were recruited online and with flyers. Inclusion criteria were 18 to 60 years of age, sufficient knowledge of German, and good health. We tested 83 people of which nine men and five women were excluded (three due to depression and two due to technical difficulties). Of the remaining 69 women, 35 were mothers (*mean* age = 39.9 years) and 34 were non-mothers (*mean* age = 36.5 years). Mothers had at least one child close to the age range of the stimuli that they encountered while non-mothers had no children of any age. Mothers and non-mothers did not differ in age, socio-economic status or intelligence (see Table 1). Participants were informed and had time to ask questions before signing a form to give their informed consent. They were reimbursed for their participation. The study was approved by the Ethics committee of the Charité – Universitätsmedizin Berlin and conducted in accordance with the Declaration of Helsinki.

Experimental procedure

Data collection consisted of an online questionnaire and in-person testing. The online questionnaire was hosted with SoSci [16]. It consisted of questions about socio-demographic status, mood status and health as well as questionnaires to measure alexithymia (TAS-20, Popp et al., 2008), emotion regulation (ERQ, Abler & Kessler, 2009) and tendency for social desirability (KSE-G, Kemper

et al., 2012). During in-person testing, they rated their mood status and completed the matrix reasoning task of the Wechsler Adult Intelligence Scale (WAIS®-IV, Wechsler, 2008), the socio-affective video task (SoVT, Klimecki et al., 2013) as well as two newly developed tasks: a morphing task to measure affect recognition and a picture-based task to measure empathy and compassion. Only the morphing task is subject of this paper. All tasks were presented on a 15-inch laptop with a resolution of 1920x1080 pixels using the software Presentation® (Neurobehavioral Systems). Participants were seated in front of the laptop with approximately 50cm distance between their eyes and the screen. Testing lasted about 60min with the affect recognition task lasting 10 to 15min.

Morphing task

We developed a new affect recognition task based on morphed videos portraying neutral child faces continually and gradually changing into an emotional face (see Fig 1). Similar morphing tasks have been successfully used before, however, using either infant or adult stimuli but never stimuli of children [6,22,23]. Faces were taken from established databases (NIMH-ChEFS, Egger et al., 2011; DEFSS, Meuwissen et al., 2017; Dartmouth Database, Dalrymple et al., 2013; CAFE Set, LoBue & Thrasher, 2015). We used 32 identities with available angry, afraid, sad, happy and neutral faces. Pictures were processed in GIMP (Version 2.8.22) to be 640x800 pixel, greyscale on a black background and only show faces without hair or ears. These pictures were morphed using WinMorph© (Version 3.01) to create 10s long videos. The 32 identities were divided into three age groups: 4 to 6 years old (pre-school), 7 to 10 years old (middle children) and 11 to 15 years old (adolescents). Each version consisted of 64 morphed videos of interest, showing 8 male and 8 female identities in each of the four emotions. Additionally, participants saw four practice videos to familiarise themselves with the task. Mothers were assigned the age group closest to their child or children while non-mothers were assigned randomly to one age group.

Fig 1. Still images from an experimental stimulus.

Still images from one of the 10s long videos in the middle age group of the affect recognition task. The neutral expression gradually and continuously changes into an angry expression. Participants are asked to stop the video as soon as they recognise the emotion and then choose the correct emotion from four options (afraid, angry, happy and sad).

Videos were presented in pseudo-randomised order so that never more than two videos in a row were of the same emotion or identity. Participants were asked to stop the video as soon as they recognised the emotion allowing response time for correct items to measure how much information a participant needed to correctly recognise the emotion. If participants did not stop the video, the response time is coded as the full duration of the video (10s). After stopping the video, participants had to choose between four options. After stopping the video, participants were asked to choose the emotion from four options (afraid, angry, happy and sad) using the arrow keys and the space bar to confirm their choice, all within 8s. However, it was also possible for them to opt out and signify that they did not recognise the target emotion. All participants used their dominant hand to press the keys. Before doing the affect recognition task, participants performed a control task where the same neutral face was morphed to animal faces using the same procedure as for morphing the neutral to emotional faces. Here, participants were asked to stop the video as soon as they recognise the species and then choose the species from the following options: ape, cat, dog and lion. This control task consisted of 16 morphed videos of interest and two practice videos.

Analysis

All analyses are based on the Bayesian framework and were performed using JASP [28]. The Bayes factor BF_{10} encodes the relative likelihood of the data given a model compared to the null model without any of the predictors. We compare all possible combinations of predictors and

interactions to the null model. In the case of multiple predictors, we also inspect the inclusion Bayes factor BF_{incl} which encodes the Bayes factor for each predictor across all matched models simultaneously. We use Jeffrey's scheme to label the strength of evidence for or against a specific model: "decisive" for a model more than 100 times more likely, "very strong" for a model between 30 and 100 times more likely, "strong" for a model between 10 and 30 times more likely, "moderate" for a model between 3 and 10 times more likely and "anecdotal" for a model between 1 and 3 times more likely. The multiplicative inverse is used for evidence for the null model [29].

Differences between mothers and non-mothers in age, socio-economic status, intelligence (WAIS®-IV - matrix reasoning task), mood status (both during online and in person testing), empathy and compassion (SoVT) as well as questionnaire scores (TAS, ERQ, KSE-G) were evaluated using Bayesian Mann-Whitney-U tests with 10,000 random samples and corrected for multiple comparisons using Westfall's method [30,31]. The socio-economic status score was computed using information on net equivalent income, education and career (SES, Lampert et al., 2013). To investigate possible differences in relationship status, a Bayesian contingency table was computed with independent multinomial sampling.

To analyse performance in the affect recognition and the control task and test our hypothesis, we used a linear integrated speed accuracy score (LISAS) combining both response times and accuracies [15]. All response times where participants stopped the video within 500ms or where they did not choose any of the four emotions were excluded from analysis. No participant lost more than 25% of trials in one task. LISAS are computed by adding the error rate multiplied with the ratio between the standard deviation of correct response times and the standard deviation of the error rate to the average correct response time. Thereby, the score integrates response times and accuracies to even out the influence of different strategies. The lower the LISAS, the better was the performance allowing analyses of differences in overall performance. Additionally, we investigated

response times and accuracy where LISAS differed to explore the origin of differences in performance.

The performance in the control task as measured by LISAS was compared between mothers and non-mothers using a Bayesian ANOVA with the predictor motherhood (yes/no). Additionally, we added the predictor age group to rule out chance differences (4-6, 7-10, 11-15). For evaluating the affect recognition task itself, Bayesian repeated measures ANOVAs were used with the following predictors: motherhood (yes or no), age group (4-6, 7-10, 11-15) and emotion (afraid, angry, sad, happy). Where the best model included the predictor age group or emotion, post hoc tests were computed to determine which levels differ. For these tests, posterior odds were corrected for multiple testing and individual comparisons are based on a t-test with a Cauchy prior [30]. Since inspection of the Q-Q plots of residuals for accuracies in the affect recognition task deviated slightly from the normal distribution, a Bayesian Mann-Whitney-U test comparing accuracies of mothers and non-mothers in the affect recognition task was computed with 10,000 random samples to confirm the result of the ANOVA.

We performed a Bayesian linear regression using the performance in the control task, emotion regulation and alexithymia to predict the performance in the affect recognition task as measured by LISAS averaged over all emotions. We used an uninformed uniform prior of 12.5% for each of the possible models.

Results

All data and relevant scripts can be found on OSF (<https://osf.io/twxsk/>).

Sample characteristics

Our samples of mothers and non-mothers were comparable in almost all accompanying measures that we investigated. They did not differ in age, intelligence, socioeconomic status, tendency to answer socially desirable, alexithymia, emotion regulation, empathy or compassion (see

Table 1). They also did not differ in their rating of their mood state before they performed the affect recognition task or when they filled out the online questionnaire. However, mothers were more likely to be in a relationship than non-mothers with 79.41% of non-mothers being single but only 44.12% of mothers ($BF_{10} = 24.69$). Therefore, mothers and non-mothers were comparable in many traits and socioeconomic factors, but they differed in relationship status.

Table 1. Group comparisons and results of the Bayesian Mann-Whitney-U tests corrected for multiple comparisons.

Measurement	mothers	non-mothers	BF_{10}	W
age	39.85 ± 1.23	36.47 ± 1.46	0.386	407
ERQ	39.12 ± 1.08	40.59 ± 1.31	0.145	709
KSE-G	1.61 ± 0.06	1.76 ± 0.06	0.303	734
mood state: online	3.94 ± 0.13	4.06 ± 0.12	0.095	604
mood state: in person	4.21 ± 0.10	4.12 ± 0.10	0.097	535
SES	13.00 ± 0.28	13.34 ± 0.47	0.157	693
SoVT compassion	63.77 ± 2.33	58.35 ± 2.88	0.168	468
SoVT empathy	44.64 ± 2.65	46.11 ± 2.92	0.082	627
TAS	37.79 ± 1.26	38.32 ± 1.15	0.087	595
WAIS®-IV (matrices)	10.58 ± 0.37	9.67 ± 0.48	0.212	417

ERQ = emotion regulation questionnaire; KSE-G = Kurzkala Soziale Erwünschtheit-Gamma (social desirability scale); SES = socioeconomic score; SoVT = socio-affective video task; TAS = Toronto alexithymia scale; WAIS®-IV = Wechsler Adult Intelligence Scale – 4th edition.

Control task

Additionally, mothers and non-mothers did not differ in their performance in the control task as measured by LISAS (see Fig 2). The Bayesian ANOVA revealed anecdotal evidence *against* the alternative model including the predictor motherhood ($BF_{10} = 0.45$), moderate evidence *against* the model including age group ($BF_{10} = 0.13$), strong evidence against the model including both predictors ($BF_{10} = 0.06$) and very strong evidence against the model including both predictors and the interaction ($BF_{10} = 0.03$). Therefore, mothers and non-mothers performed equally well in a comparable recognition task not involving emotions and there were no differences between the participants assigned to the different age groups for the affect recognition task. This indicates that any differences that are found between the groups in the affect recognition task cannot be

attributed to overall differences in reaction speed, concentration or other processes that are not related with processing and recognising facial affect.

Fig 2. LISAS in the control task

Performance in the control task separately for mothers and non-mothers as well as age groups in the affect recognition task. As confirmed by the Bayesian ANOVA, the best model describing the data is the null model suggesting that mothers and non-mothers and all groups performed equally well.

Affect recognition task

The Bayesian repeated measures ANOVA looking into the effect of motherhood (yes or no), emotion (afraid, angry, happy and sad), age group (4-6, 7-10, 11-15 years old) and their interactions on LISAS of the affect recognition task revealed decisive evidence in favour of the model involving all three predictors and the interaction between age group and emotion (see Table 2 and Fig 3). This model was more than four times as likely as the second-best model which additionally included the interaction between motherhood and age group. This is also reflected in the analysis of effects over matched models with decisive evidence for the inclusion of the predictor emotion ($BF_{incl} = 8.81e+31$) and the interaction between emotion and age group ($BF_{incl} = 6.96e+3$), strong evidence in favour of the predictor motherhood ($BF_{incl} = 18.16$) and moderate evidence in favour of the predictor age group ($BF_{incl} = 6.73$). There was strong evidence against the inclusion of the three-way interaction ($BF_{incl} = 0.08$), moderate evidence against the interaction between motherhood and emotion ($BF_{incl} = 0.17$) and the interaction between motherhood and age group ($BF_{incl} = 0.23$).

245 *Table 2. Results of the Bayesian repeated measures ANOVA investigating differences in LISAS in the affect recognition*
 246 *task.*

Models	$P(M data)$	BF_{10}	error %
Null model	2.17e-38	1.000	
emotion + motherhood + age group + emotion * age group	0.66	3.04e+37	3.02
emotion + motherhood + age group + emotion * age group + motherhood * age group	0.16	7.33e+36	7.32
emotion + motherhood + age group + emotion * motherhood + emotion * age group	0.13	5.74e+36	33.94
emotion + age group + emotion * age group	0.04	1.67e+36	0.93
emotion + motherhood + age group + emotion * motherhood + emotion * age group + motherhood * age group	0.02	8.17e+35	10.58
emotion + motherhood + age group + emotion * motherhood + emotion * age group + motherhood * age group + emotion * motherhood * age group	0.00	6.28e+34	23.49
emotion + motherhood + age group	1.04e-4	4.76e+33	4.20
emotion + motherhood + age group + motherhood * age group	2.12e-5	9.75e+32	5.39
emotion + motherhood	1.53e-5	7.05e+32	7.82
emotion + motherhood + age group + emotion * motherhood	1.09e-5	5.03e+32	1.68
emotion + age group	5.58e-6	2.57e+32	3.64
emotion + motherhood + age group + emotion * motherhood + motherhood * age group	2.20e-6	1.01e+32	5.55
emotion + motherhood + emotion * motherhood	1.55e-6	7.11e+31	1.43
emotion	9.60e-7	4.41e+31	0.99
motherhood + age group	1.14e-36	52.31	0.93
motherhood	2.53e-37	11.65	0.94
motherhood + age group + motherhood * age group	1.58e-37	7.3	5.03
age group	9.40e-38	4.33	1.20

Fig 3. LISAS per emotion and age group, separately for mothers and non-mothers.

As confirmed by the Bayesian repeated measures ANOVA, the best model describing the data included the predictors emotion, motherhood, age group as well as the interaction between age group and emotion. Mothers performed worse overall and both mothers and non-mothers performed worst in the youngest age group. There were also differences in performance due to the different emotions with happy being the easiest emotion to recognise.

Contrary to our hypothesis, mothers performed worse instead of better across all age groups and emotions than non-mothers. Post-hoc comparisons showed that both mothers and non-mothers performed worse in the youngest age group compared to the oldest age group confirming our hypothesis that performance is higher in response to adolescent faces ($BF_{10,U} = 373.25$). There is also anecdotal evidence that they performed worse in the youngest compared to the middle group ($BF_{10,U} = 1.17$), however, there is anecdotal evidence *against* differences in performance between

the middle age group and the oldest age group ($BF_{10,U} = 0.92$). Additionally, both mothers and non-mothers performed best when recognising happiness (happy versus afraid: $BF_{10,U} = 7.09e+17$; happy versus angry: $BF_{10,U} = 4.41e+9$; happy versus sad: $BF_{10,U} = 2.93e+15$) and worst when recognising sadness or fear (afraid versus angry: $BF_{10,U} = 9.58$; afraid versus sad: $BF_{10,U} = 2.44$; angry versus sad: $BF_{10,U} = 5.68e+8$). The effect of age group interacted with the effect of emotion with differences between age groups seeming to be the most pronounced in happy faces followed by afraid faces (see Table 2).

To explore the origin of the differences in performance, the same predictors were entered into two additional Bayesian repeated measures ANOVAs, one for response time and one for accuracies (see Table S1 and S2 in the supplementary materials). Investigating the response times of correct items more closely revealed the same effects as the analysis of the LISAS. There was decisive evidence in favour of the model including motherhood, emotion, age group and the interaction of the last two compared with the null model ($BF_{10} = 1.97e+33$). This model performed three times as well as the second-best model which additionally included the interaction between motherhood and age group ($BF_{10} = 6.46e+32$). Analysis of effects again revealed decisive evidence for the inclusion of the predictor emotion ($BF_{incl} = 2.33e+26$) and the interaction between age group and emotion ($BF_{incl} = 3.07e+5$) as well as strong evidence for the predictor motherhood ($BF_{incl} = 14.89$) and anecdotal evidence for age group ($BF_{incl} = 2.57$). There was strong evidence against the inclusion of the three-way interaction ($BF_{incl} = 0.04$), moderate evidence against the interaction between emotion and motherhood ($BF_{incl} = 0.14$) and the interaction between age group and motherhood ($BF_{incl} = 0.32$). Mothers had higher response times than non-mothers in all emotions and age groups, which adds to the evidence against our hypothesis. Post hoc tests again revealed longer reaction times in mothers and non-mothers in response to the youngest age group compared to the oldest but indicated no difference between the middle group and the oldest or youngest group (see Table S1 in the supplementary materials). Happiness was recognised fastest by both mothers and non-mothers

followed by anger, then fear and then sadness. Differences between age groups again seem largest in happiness followed by fear (see Table 3).

Table 3. Average and standard error of response times (RT) for correct items and accuracies in the affect recognition task for age groups, motherhood and emotions separately.

	4-6		7-10		11-15	
	RT	Accuracy	RT	Accuracy	RT	Accuracy
afraid						
mothers	7350 ± 258	73.8 ± 4.7	7085 ± 487	75.7 ± 4.0	6401 ± 544	88.8 ± 3.5
non-mothers	6505 ± 267	74.5 ± 5.3	6165 ± 240	77.3 ± 5.1	5440 ± 443	90.0 ± 3.1
angry						
mothers	6343 ± 204	76.2 ± 2.2	6251 ± 450	89.6 ± 2.8	6508 ± 445	80.6 ± 4.5
non-mothers	5923 ± 164	78.8 ± 5.1	5841 ± 200	84.7 ± 3.2	5795 ± 343	83.8 ± 3.0
happy						
mothers	6128 ± 283	93.4 ± 2.0	5858 ± 436	94.4 ± 2.8	4563 ± 565	97.5 ± 1.4
non-mothers	5601 ± 262	94.2 ± 1.3	4848 ± 514	95.5 ± 2.2	3615 ± 319	96.9 ± 1.0
sad						
mothers	7374 ± 197	77.3 ± 3.4	7098 ± 515	68.8 ± 4.7	7150 ± 414	75.0 ± 5.7
non-mothers	6747 ± 206	78.8 ± 4.1	6405 ± 187	64.2 ± 5.4	6280 ± 268	56.9 ± 4.5

Investigating accuracies, the model that described the data best included the predictors emotion, age group and their interaction but did not include the predictor motherhood ($BF_{10} = 9.72e+21$). There is decisive evidence in favour of this model compared to the null model and it performs more than five times as well as the second-best model which additionally includes the predictor motherhood ($BF_{10} = 1.90e+21$). Analysis of the inclusion of the effects revealed decisive evidence for the predictor emotion ($BF_{incl} = 3.75e+18$) and the interaction between emotion and age group ($BF_{incl} = 2.35e+4$). There was moderate evidence against the inclusion of any other predictor or interaction (age group $BF_{incl} = 0.12$; motherhood $BF_{incl} = 0.20$; emotion * motherhood $BF_{incl} = 0.21$; age group * motherhood $BF_{incl} = 0.17$; emotion * motherhood * age group $BF_{incl} = 0.22$). Post hoc tests did not show any significant differences between the age groups (see Table S2 in the supplementary materials). In terms of emotions, happy faces were recognised with the highest accuracy by both mothers and non-mothers, followed by afraid and angry faces which did not differ in accuracy. Sad faces were recognised the least accurately. Differences between age groups seemed

the most pronounced in afraid and sad faces and the least pronounced in happy faces, possibly due to a ceiling effect.

Visual inspection of the Q-Q plot of the residuals revealed slight deviations from normal distribution. Therefore, a Mann-Whitney-U test was computed with 10,000 random samples to investigate the influence of the predictor motherhood on the average accuracy over all emotions and age groups. It revealed moderate evidence against the predictor motherhood confirming the results from the ANOVA ($BF_{10} = 0.31$, $W = 537$). Mothers therefore did not differ from non-mothers in the accuracy with which they recognised the emotions. Taken together with the results of the ANOVA investigating response times, this indicates that the decrease in performance of mothers compared to the performance of non-mothers was due to the mothers needing more information (i.e., longer response times) to recognise the emotion with the same accuracy.

Predicting affect recognition performance

A Bayesian linear regression revealed that the performance in the affect recognition task was only predicted by the performance in the control task with decisive evidence for this model as compared to the null model (see Table 4). This model was more than twice as likely as the second-best model which additionally included the predictor emotion regulation. The posterior summary confirms this conclusion with very strong evidence in favour of the inclusion of the performance in the control task ($BF_{incl} = 9.75e+2$) and anecdotal evidence against the predictor alexithymia ($BF_{incl} = 0.50$) and emotion regulation ($BF_{incl} = 0.94$). This indicates that the performance in the affect recognition task is not predicted by emotion regulation or alexithymia but can be predicted by the performance in the control task. To derive the regression equation in Equation 1, a simple linear regression was performed with the LISAS in the control task as predictor for LISAS in the affect recognition task. This adds further evidence that the control task relies on similar processes without including affect processing or affect recognition.

Equation 1. Regression equation to predict LISAS in the affect recognition task using LISAS in the control task.

$$LISAS(affect) = 6853.006 + 0.536 * (LISAS(control) - 4795.091)$$

Table 4. Model comparison in the multiple linear regression to predict LISAS in the affect recognition task using LISAS in the control task, emotion regulation and alexithymia.

Models	$P(M data)$	BF_M	BF_{10}	R^2
Null model	7.98e-4	0.00	1.00	0.00
LISAS control task	0.41	7.55	1.53e+3	0.26
LISAS control task + emotion regulation	0.26	3.80	9.66e+2	0.28
LISAS control task + alexithymia	0.11	1.32	4.03e+2	0.26
LISAS control task + emotion regulation + alexithymia	0.23	0.89	2.86e+2	0.28
Alexithymia	1.14e-4	0.00	0.43	0.02
Emotion regulation	7.13e-5	7.85e-4	0.27	0.00
Emotion regulation + alexithymia	4.16e-5	4.58e-4	0.16	0.02

Discussion

This study shows a difference between mothers and non-mothers in affect recognition of an unfamiliar child. Contrary to our hypothesis, mothers exhibited a lower performance than non-mothers, needing more information to recognise an emotion with the same accuracy as non-mothers. This effect persists equally in the three age groups and is independent of emotion. No matter whether a child displays anger, sadness, fear or happiness, mothers take longer to correctly identify the emotion with the same accuracy as non-mothers. Additionally, it was easier for both mothers and non-mothers to recognise the emotion in adolescent faces than in pre-school child faces as reflected in the integrated performance scores. This effect of age group was independent of motherhood but interacted with the emotion displayed with more pronounced differences between age groups in happy and afraid faces. Therefore, while there were effects of age group and of motherhood, there was no interaction between age group and motherhood. Furthermore, the effect of motherhood was inverse to our expectations with mothers needing more information than non-mothers to reach the same level of accuracy.

Mothers had a lower performance as measured by the LISAS than non-mothers in the affect recognition task independent of age group and emotion which is at odds with our hypothesis. Our analyses revealed that this lower performance was due to mothers needing more time and therefore more information to recognise emotions with the same accuracy as non-mothers. A similar effect has been reported in a study primarily investigating effects of parenthood and gender on event-related potentials (ERPs; Proverbio et al., 2006). In this study, participants saw infant faces and were asked to judge as accurately and quickly as possible whether the emotional state displayed was positive or negative. Parents were slower than non-parents in this task. This was accompanied by differences in several ERPs, most notably in the N245 and the P3. The N245 has been associated with sensitivity to emotional faces, indicating that parents were more sensitive to the emotions in the infant faces [34]. Additionally, the authors suggest that the increased P3 in mothers may have the function to induce empathic behaviour and that the slower response time may have been due to the greater emotional impact of the infant faces on parents. Our study therefore replicates Proverbio's findings and extends them to child faces in multiple age groups expressing multiple emotions showing that the effect of motherhood is independent of emotion and age group.

Many studies on social understanding indicate that mothers might be more responsive to emotional child faces than non-mothers. Mothers experience more distraction by emotional infant and pre-pubescent faces when performing an unrelated task [13,35] and rate the valence of infant faces as more extreme [36]. These differences also extend to the neural level with mothers showing a stronger neural empathy response to both adult and child stimuli [37]. In an affective theory of mind task, Plank and colleagues found that mothers exhibited stronger activation in both emotion processing and theory of mind areas than non-mothers in response to both affective adult and affective child faces [8]. These studies indicate that mothers seem to respond more strongly to affective stimuli than non-mothers. This increased emotional response could be the reason for the delayed response in this study. Since parenting has a strong affective component, this increased

emotional response might increase mothers' affective social understanding and react with more compassion to other's in distress.

Our samples of mothers and non-mothers performed equally well in a control task where they were asked to recognise animals. This control task was designed to be highly similar to the affect recognition task but did not include any emotional stimuli. A neutral human face continually and gradually changed into an animal's face. The performance in both tasks, therefore, depended on general attentive processes, reaction speed, concentration, visual processing and face perception. However, only the affect recognition task included the processing and recognition of facial affect. Mothers and non-mothers performed equally well on the control task, indicating that they are equally capable in detecting these continuous and gradual changes in faces. Therefore, we can rule out that differences in the affect recognition task between mothers and non-mothers are due to differences in general cognitive capacity or task competency.

Contrary to our hypothesis, there was no interaction between the effect of motherhood and age group. This indicates that the effect of motherhood on affect recognition is stable over childhood and adolescence. It also suggests that even when features become increasingly adult, the difference in performance between mothers and non-mothers persists. This could be due to the prolonged period of dependence in humans. With parenthood extending into adolescence and early adulthood of the offspring, differences persisting into these age groups are to be expected. It is possible that this difference between mothers and non-mothers extends to adult faces or that it vanishes for young or older adults.

Emotions in faces of adolescents were easier to recognise for both mothers and non-mothers than faces of younger children. This ease of affect recognition could be due to adolescents decreased Kindchenschema and increasingly adult features. Similarly to Thompson-Booth and colleagues data (2014), pre-school children were treated differently as compared to adolescents. The group of 7- to 10-year-old children are midway between pre-school children and adolescents

with no significant difference in performance to either of those groups. This is in line with previous research on attentional allocation [13]. It is possible that these differences between age groups are not due to differences in the way different age groups express a specific emotion but in the effect the stimulus has on the perceiver with emotional child faces leading to task interference and therefore decreased performance. Studies incorporating multiple tasks focussing on different processes of face and affect processing could help disentangle these possibilities.

Mothers and non-mothers were well-matched in age, socio-economic status, emotion regulation, intelligence and alexithymia. One aspect where mothers and non-mothers differed was relationship status with mothers being more likely to be in a relationship than non-mothers. It is possible that affect recognition is influenced by relationship status. Close relationships entail reading various emotions in another person's face, while less close or more professional relationships are often confined in emotional variability and frequency. Therefore, we cannot rule out that relationship status affected our results based on this sample. Studies with a balanced sample are needed to investigate the interaction of parenthood and relationship status on affect recognition.

In the here presented study, mothers differed from matched non-mothers in their performance in a facial affect recognition task with stimuli of unfamiliar children. This difference in performance was due to the mothers needing more time and therefore more information to achieve the same accuracy as non-mothers. Both groups performed equally well in a control task that required them to recognise animals instead of emotions. The effect of motherhood was independent of the age group of the stimuli and of the expressed emotion. These results contradict the hypothesis that motherhood leads to better performance in recognising emotions in child faces. The decreased performance in mothers might be due to their increased emotional response to children in general and affective child faces in particular [8,37]. Additionally to a clear effect of motherhood, we found an effect of age group. Adolescent stimuli were easier to interpret for both mothers and non-mothers than pre-school children. A possible explanation is their increasingly adult features as well

as their decreased Kindchenschema. This effect interacted with the emotion expressed, with the strongest difference in happy followed by afraid faces. Further research is needed to examine possible causes of the decreased performance in mothers including differences in emotional response as well as whether the effect extends to adult faces.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF; grant 01KR1803C; <https://www.ubica.site>) and by the “Dissertation Plus” program of the Claussen-Simon-Stiftung.

Competing interests

The authors declare no competing interests.

433 References

- 434 1. Pereira M, Ferreira A. Neuroanatomical and neurochemical basis of parenting: dynamic
435 coordination of motivational, affective and cognitive processes. *Horm Behav.* 2016;77: 72–85.
436 doi:10.1016/j.yhbeh.2015.08.005
- 437 2. Swain JE, Kim P, Spicer J, Ho SS, Dayton CJ, Elmadih A, et al. Approaching the biology of
438 human parental attachment: Brain imaging, oxytocin and coordinated assessments of
439 mothers and fathers. *Brain Res.* 2014;1580: 78–101. doi:10.1016/j.brainres.2014.03.007
- 440 3. Decety J. The neuroevolution of empathy. *Ann N Y Acad Sci.* 2011;1231: 35–45.
441 doi:10.1111/j.1749-6632.2011.06027.x
- 442 4. Feldman R. The adaptive human parental brain: Implications for children’s social
443 development. *Trends Neurosci.* 2015;38: 387–399. doi:10.1016/j.tins.2015.04.004
- 444 5. Feldman R. Oxytocin and social affiliation in humans. *Horm Behav.* 2012;61: 380–391.
445 doi:10.1016/j.yhbeh.2012.01.008
- 446 6. Matsunaga M, Tanaka Y, Myowa M. Maternal nurturing experience affects the perception
447 and recognition of adult and infant facial expressions. *PLoS One.* 2018;13: 1–17.
448 doi:10.1371/journal.pone.0205738
- 449 7. Nishitani S, Doi H, Koyama A, Shinohara K. Differential prefrontal response to infant facial
450 emotions in mothers compared with non-mothers. *Neurosci Res.* 2011;70: 183–188.
451 doi:10.1016/j.neures.2011.02.007
- 452 8. Plank IS, Hindi Attar C, Kunas SL, Dziobek I, Bermpohl F. Motherhood and theory of mind:
453 increased activation in affective theory of mind and emotion processing areas. *OSF Prepr.*
454 2021. Available: osf.io/2cxks
- 455 9. Sullivan MW, Lewis M. Emotional expressions of young infants and children a practitioner’s

- 456 primer. *Infants Young Child*. 2003;16: 120–142. doi:10.1097/00001163-200304000-00005
- 457 10. Lucion MK, Oliveira V, Bizarro L, Bischoff AR, Silveira PP, Kauer-Sant’Anna M. Attentional bias
458 toward infant faces – Review of the adaptive and clinical relevance. *Int J Psychophysiol*.
459 2017;114: 1–8.
- 460 11. Kringelbach ML, Stark EA, Alexander C, Bornstein MH, Stein A. On cuteness: unlocking the
461 parental brain and beyond. *Trends Cogn Sci*. 2017;20: 545–558.
462 doi:10.1016/j.tics.2016.05.003
- 463 12. Luo L, Li H, Lee K. Are children’s faces really more appealing than those of adults? Testing the
464 baby schema hypothesis beyond infancy. *J Exp Child Psychol*. 2011;110: 115–124.
465 doi:10.1016/j.jecp.2011.04.002
- 466 13. Thompson-Booth C, Viding E, Mayes LC, Rutherford HJV, Hodson S, McCrory EJ. I can’t take
467 my eyes off of you: Attentional allocation to infant, child, adolescent and adult faces in
468 mothers and non-mothers. *PLoS One*. 2014;9: e109362.
- 469 14. BMFSFJ. Gender Care Gap - ein Indikator für die Gleichstellung [Accessed: 2020-12-16].
470 Bundesministerium für Familie, Senioren, Frauen und Jugend; 2019. Available:
471 [https://www.bmfsfj.de/bmfsfj/themen/gleichstellung/gender-care-gap/indikator-fuer-die-](https://www.bmfsfj.de/bmfsfj/themen/gleichstellung/gender-care-gap/indikator-fuer-die-gleichstellung/gender-care-gap---ein-indikator-fuer-die-gleichstellung/137294)
472 [gleichstellung/gender-care-gap---ein-indikator-fuer-die-gleichstellung/137294](https://www.bmfsfj.de/bmfsfj/themen/gleichstellung/gender-care-gap/indikator-fuer-die-gleichstellung/137294)
- 473 15. Vandierendonck A. Further Tests of the Utility of Integrated Speed-Accuracy Measures in Task
474 Switching. *J Cogn*. 2018;1: 1–16.
- 475 16. Leiner DJ. SoSci Survey (Version 3.1.06) [Computer software]. 2019. Available:
476 <https://www.soscisurvey.de>
- 477 17. Popp K, Schäfer R, Schneider C, Brähler E, Decker O, Hardt J, et al. Faktorstruktur and
478 Reliabilität der Toronto-Alexithymie-Skala (TAS-20) in der deutschen Bevölkerung. *Psychother*
479 *Psychosom Medizinische Psychol*. 2008;58: 208–214. doi:10.1055/s

- 480 18. Abler B, Kessler H. Emotion Regulation Questionnaire - Eine Deutschsprachige Fassung des
481 ERQ von Gross und John. *Diagnostica*. 2009;55: 144–152.
- 482 19. Kemper C, Beierlein C, Bensch D, Kovaleva A, Rammstedt B. Eine Kurzsкала zur Erfassung des
483 Gamma-Faktors sozial erwünschten Antwortverhaltens: Die Kurzsкала Soziale Erwünschtheit-
484 Gamma (KSE-G). *GESIS - Work Pap.* 2012;25.
- 485 20. Wechsler D. Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV). APA PsycTests;
486 2008. Available: <https://psycnet.apa.org/doi/10.1037/t15169-000>
- 487 21. Klimecki OM, Leiberg S, Lamm C, Singer T. Functional neural plasticity and associated changes
488 in positive affect after compassion training. *Cereb Cortex*. 2013;23: 1552–1561.
- 489 22. Domes G, Czeschnek D, Weidler F, Berger C, Fast K, Herpertz SC. Recognition of facial affect
490 in borderline personality disorder. *J Pers Disord*. 2008;22: 135–147.
- 491 23. Schwenck C, Mergenthaler J, Keller K, Zech J, Salehi S, Taurines R, et al. Empathy in children
492 with autism and conduct disorder: Group-specific profiles and developmental aspects. *J Child*
493 *Psychol Psychiatry*. 2012;53: 651–659.
- 494 24. Egger HL, Pine DS, Nelson E, Leibenluft E, Ernst M, Towbin KE, et al. The NIMH Child
495 Emotional Faces Picture Set (NIMH-ChEFS): a new set of children’s facial emotion stimuli. *Int J*
496 *Methods Psychiatr Res*. 2011;20: 145–156. doi:10.1002/mpr.343
- 497 25. Meuwissen AS, Anderson JE, Zelazo PD. The creation and validation of the Developmental
498 Emotional Faces Stimulus Set. *Behav Res Methods*. 2017;49: 960–966. doi:10.3758/s13428-
499 016-0756-7
- 500 26. Dalrymple KA, Gomez J, Duchaine B. The dartmouth database of children’s faces: Acquisition
501 and validation of a new face stimulus set. *PLoS One*. 2013;8: 1–7.
- 502 27. LoBue V, Thrasher C. The Child Affective Facial Expression (CAFE) set: Validity and reliability

503 from untrained adults. *Front Psychol.* 2015;5: 1–8. doi:10.3389/fpsyg.2014.01532

504 28. JASP Team. JASP (Version 0.14.1)[Computer software]. 2020. Available: <https://jasp-stats.org/>

505 29. Goss-Sampson M. Bayesian Inference in JASP. 2020. pp. 1–120. doi:10.17605/OSF.IO/CKNXM

506 30. Westfall PH, Johnson WO, Utts JM. A Bayesian perspective on the Bonferroni adjustment.
507 *Biometrika.* 1997;84: 419–427. doi:10.1093/biomet/84.2.419

508 31. de Jong T. A Bayesian Approach to the Correction for Multiplicity. *psyArXiv*; 2019.
509 doi:10.31234/osf.io/s56mk

510 32. Lampert T, Kroll LE, Müters S, Stolzenberg H. Messung des sozioökonomischen Status in der
511 Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). *Bundesgesundheitsblatt -*
512 *Gesundheitsforsch - Gesundheitsschutz.* 2013;56: 631–636. doi:10.1007/s00103-012-1663-4

513 33. Proverbio AM, Brignone V, Matarazzo S, Del Zotto M, Zani A. Gender and parental status
514 affect the visual cortical response to infant facial expression. *Neuropsychologia.* 2006;44:
515 2987–2999. doi:10.1016/j.neuropsychologia.2006.06.015

516 34. Streit M, Wölwer W, Brinkmeyer J, Ihl R, Gaebel W. Electrophysiological correlates of
517 emotional and structural face processing in humans. *Neurosci Lett.* 2000;278: 13–16.
518 doi:10.1016/S0304-3940(99)00884-8

519 35. Thompson-Booth C, Viding E, Mayes LC, Rutherford HJV, Hodsoll S, McCrory EJ. Here’s looking
520 at you, kid: Attention to infant emotional faces in mothers and non-mothers. *Dev Sci.*
521 2014;17: 35–46.

522 36. Parsons CE, Young KS, Jegindoe Elmholdt EM, Stein A, Kringelbach ML. Interpreting infant
523 emotional expressions: Parenthood has differential effects on men and women. *Q J Exp*
524 *Psychol.* 2016;70: 554–564. doi:10.1080/17470218.2016.1141967

525 37. Plank IS, Hindi Attar C, Kunas S, Dziobek I, Bermpohl F. Motherhood and Empathy: Increased

526 Activation in Empathy Areas in Response to Other's in Pain. ResearchSquare. 2021.
527 doi:10.21203/rs.3.rs-478801/v1

528

529 Supporting information

530 **S1 Table. Results of the Bayesian repeated measures ANOVA investigating differences in response**
531 **times of the correct items in the affect recognition task.**

532 **S2 Table. Results of the Bayesian repeated measures ANOVA investigating differences in**
533 **accuracies in the affect recognition task.**