Dynamics of Non-verbal Vocalizations and Hormones during Father-Infant Interaction

Omri Weisman, Mohamed Chetouani, Catherine Saint-Georges, Nadège Bourvis, Emilie Delaherche,
Orna Zagoory-Sharon, David Cohen, Ruth Feldman

Abstract—Although researchers have established the roles of oxytocin (OT) in promoting affiliative bonds and cortisol (CT) in adapting to stress, the investigation of their interplay with non-verbal behaviors has only recently begun. In this study, we employed social signal-processing techniques to investigate relationships between non-verbal features: infant and father vocalizations, infant-directed speech, speech turn-taking (STT) and hormonal dynamics (OT and CT). Thirty-five fathers were asked to interact with their infants following the fathers self-administration of OT or placebo. We consider the three episodes of the Still Face (SF) paradigm: (1) a baseline normal interaction episode, (2) the SF episode, in which the father becomes unresponsive and maintains a neutral facial expression, and (3) a reunion in which parents and their infants re-engage in interaction. This paradigm elicited stress in the infant. Statistical relationships are assessed by correlation analysis and linear mixed models (LMMs). The results indicate that (i) infant vocalization and STT are key social cues regulating interactions during the stress-inducing and reunion episodes, with infant vocalization leading the interaction dynamics; (ii) father empty pause was the main adaptive behavior of fathers after SF; (iii) OT did not modulate infant STT or father STT/fatherese; (iv) CT appeared to modulate the interaction.

Index Terms—Social Signal Processing, Parent-Infant Interaction, Hormonal modulation, Inter-personal Synchrony

◆

1 INTRODUCTION

Currently, in social signal processing and affective computing, manual annotation, self-reports and questionnaires play a major role in both the design and evaluation of computational models. The design of machines capable of perceiving and understanding human behaviors usually exploits automatic mapping of non-verbal features (e.g., facial expressions, tone of the voice, etc.) into psychometric measurements that are collected by these methodologies. Most of the previous research on these domains has focused on the processing of basic emotional states (e.g., the six basic emotions) [1]; nevertheless, there has recently been a shift toward continuous dimensional constructs, such as arousal-valence [2].

However, training computational models with such manual annotations is complex, and several methods have been investigated to overcome this complexity, including crowdsourcing [3]. Another recently proposed option in multimedia indexing is the exploitation of behaviors displayed by users while perceiving images and/or videos to automatically tag these materials (e.g., instead of asking users to use words) [4].

In this paper, we introduce a radically new framework that uses biology as the ground truth (here, hormones) to characterize social interactions. This paper can be considered as a first step towards the realization of this framework. Specifically, we investigate the interplay between non-verbal features and hormonal changes during parent-infant interaction through correlation analysis.

The next section introduces the hormones that have been studied, namely, oxytocin (OT) and cortisol (CT). We also report the results of other studies of the impacts of these hormones on behaviors and parental bonding. The aims and hypotheses of this study are reported in section 3. Section 4 describes the behavioral paradigm that was employedparent-infant interactionand the method used to measure the hormonal changes. Next, we present the automatically extracted non-verbal cues that were used to characterize the interaction dynamics in section 5. In section 6, we present the observed correlations between hormonal changes and these non-verbal cues. Section 7 summarizes and discusses our results, and we conclude in section 8.

- *O. Weisman is with the Child Study Center, Yale School of Medicine, New Haven, CT 06519, USA (e-mail: omriwei@yahoo.com)*
- *M. Chetouani and E. Delaherche are with the Institute for Intelligent Systems and Robotics, Université Pierre et Marie Curie- Paris6, 75005 Paris, France (e-mail: mohamed.chetouani@upmc.fr)*
- *C. Saint-Georges and D. Cohen are with the Department of Child and Adolescent Psychiatry, Hôpital de la Pitie-Salpêtriere, University Pierre and Marie Curie, 75013 Paris, France, and also with the Institute of Intelligent Systems and Robotics, University Pierre and Marie Curie, 75005 Paris, France (e-mail:catherine.saintgeorges@gmail.com; david.cohen@psl.aphp.fr).*
- *O. Zagoory-Sharon and R. Feldman are with the Department of Psychology, Bar-Ilan University, Ramat Gan 52900, Israel (e-mail: feldman@mail.biu.ac.il).*

2 SOCIAL BONDING AND STRESS

2.1 Nature of social cues

Several behavioral and physiological measures have been used as indices of social bonding and stress. In animal studies, depending on species and individuals, observations have shed light on a variety of social cues involved in bonding and affiliative behaviors. These studies identify the source (infant vs. caregiver), modality (e.g., audio modality) and direction (infant towards caregiver vs. caregiver towards infant) (see Table 1) of this bonding [5]. Intra-species social cues are concordant, but their meaning and relevance for each species may be different [6].

In small-brained mammals, social recognition and bonding are primarily dependent on the olfactory system, whereas in primates, the role of olfactory cues is relatively limited. In humans, infants seem to specifically exploit the audio modality to inform their caregiver of their needs (e.g., an infant may express herself by crying). Meanwhile, because caregivers are sensitive to this modality, they also exploit vocalization to regulate interactions. This type of vocalization is called parentese or infant-directed speech [7] and can be specified as motherese or fatherese depending on the sex of the parent.

2.2 Hormonal basis of bonding and stress

These social cues play an important role in facilitating maternal care and ensuring mother-infant bond formation and stress adaptation. The biological basis of these bonding and adaptability phenomena has also been investigated. Several studies have shown that OT plays a fundamental role in establishing social and caregiver-infant bonds in animals [5] and humans [19], [20]. OT is known to be implicated in mammalian sociality [21]. Conceptual models addressing the mechanisms underlying the effects of OT on social processes suggest two types of mechanisms. The first one is the Social Salience hypothesis, which suggests that OT increases the salience of social cues. The second one is the Reward hypothesis, which postulates that OT increases the brain's reward response from social stimuli. Both mechanisms may be active in the context of the brain's response under the effect of OT [22]. Interestingly, OT modulates the strength of the social bond, social emotion and behaviors, such as trust, generosity, altruism, fidelity, empathy, and social memory [23]. Regarding early infant/caregiver interaction, OT levels in pregnant women predict the expression of maternal behaviors in the post-partum period, suggesting a priming effect of OT on the initiation of parenting behavior [16], [24].

CT is the most researched biomarker of the stress response, and its release is highly sensitive to the social context [25]. Extensive research has revealed increases in CT, especially salivary CT, during moments of stress that involve social evaluation, public speaking, stressful interactions, or encountering fear- or anxiety-promoting experiences (for a review, [26], [27]). Additionally, studies have shown associations between activations of brain circuits related to vigilance and fear (e.g., the amygdala network) with elevations in CT response (for a review, [28]). In the context of early infant/caregiver interaction, a synchronous and positive parent-child relationship is associated with a lower baseline CT [29], while psychopathological conditions, such as maternal anxiety and depression, correlate with higher CT responses [30].

The associations between OT and CT are also of interest. Several studies have revealed that OT administration decreases CT production, for instance, in the context of couple conflict discussion [31] or social stress paradigms [32]. Additionally, in the context of family interactions, OT and CT have been found to each be independently predictive of a positive mutual atmosphere within the family [33]. These findings highlight the balance between the stress and affiliation systems, particularly between OT and CT, and their online interaction as important neuroendocrine underpinnings of the individual's physiological response to social processes and stress.

2.3 The Still Face (SF) paradigm: promoting bonding and eliciting stress

The developmental effects of caregiver-infant bonding have also been studied through a specific experimental paradigm created to model consequences of parental early deprivation: the Still Face (SF) paradigm.

The SF paradigm considers three different phases: (1) a baseline normal interaction episode, (2) the SF episode, in which the parents become unresponsive and maintains a neutral facial expression, and (3) a reunion in which parents and their infants re-engage in interaction. Over the past three decades, Tronick's observational paradigm [34] has been used in numerous studies to understand infant socio-emotional regulation strategies and stress coping.

The SF paradigm allowed characterizing the dynamics of the interaction (also known as dyadic flexibility) and the change (or stability) in infant-caregiver dyads. SF is a stressful condition for the infant because the caregiver suddenly interrupts the interaction and remains still for a given period of time [35]. This stressful social situation appears to amplify individual differences in emotionality [36], [34].

In the reunion episode, parents and their infants re-engage in interaction. This episode usually highlights individual differences in dyadic coordination and parental support of infants' emotion regulation [37].

3 STUDY AIMS AND HYPOTHESES

This paper aims to demonstrate that automatic behavior analysis methods rooted in social signal processing

TABLE 1
Social cues involved in bonding and affiliative behavior during infant caregiver early interaction

Modality	Type of cues	Species	Reference
From Infant to Caregiver			
Audio	Ultrasonic vocalization	Mice	Branchi, 2001 [8]
	Crying	Human	Falk, 2004 [9]
Olfactory	Pup odor	Mice	Levy et al., 2004 [10]
Visual	Gazing / eye contact	Primate / human	Bjorklund, 1987 [11]
Physical	Stroking	Primate / human	Franceschini, 1989 [12]
From Caregiver to Infant			
Audio	Motherese	Human	Falk, 2004 [9]
Olfactory	Nipple odor	All mammals	Moriceau, 2004 [13]
Visual	Motionese	Bird	Lorenz, 1935 [14]
	Imitating	Human	Paukner, 2011 [15]
Physical	Holding / handling	Primate / human	Falk, 2004 [9]
	Affectionate touch	Primate / human	Feldman, 2007 [16]
Social	Synchronizing	Human	Fedlman, 2003 [17]
	Shared parenthood	Human	Rilling, 2013 [18]

and affective computing may provide insights into the interplay between social signals and hormonal changes during the SF paradigm.

Figure 1 details the main steps of our approach. Experiments were designed by psychologists by considering two parameters that should potentially modulate these social signals: treatment condition (OT vs. placebo) and time condition (before vs. after SF). This design did not initially involve any computational dimension, and consequently, the audiovisual conditions were challenging (e.g., moving cameras/microphones and/or different resolutions). Given these uncontrolled situations, we opted for a manual segmentation of both infant and father vocalizations and the annotation of infant-directed speech events (fatherese) (see section 5.1). This segmentation follows a machine learning approach for the automatic detection of fatherese. Then, individual and dyadic features are extracted to capture the dynamics of social exchanges (see section 5.2). Finally, these features are correlated with hormonal changes (here, OT and CT) (see section 6). LMMs are employed to assess the interplay between experimental conditions (OT, placebo, SF conditions, and visits), non-verbal features and hormones.

This computational approach allows us to investigate the following research questions: (1) OT dynamics and parents' vocalizations. In children aged 7-12 years, Seltzer et al. [38] showed that mothers' vocalizations contributed to the OT regulation of social bonding and adaptation to stressors. The automatic characterization of parents' vocalizations is investigated for evaluating the influence of OT in the case of parent-infant interactions. (2) Does baseline CT or a change in CT during SF alter the quality of the interaction? (3) Are infant or caregiver vocalizations the only variables to consider? Should we consider Speech Turn-Taking (STT) and dynamics as well? Finally, how are the interaction dynamics expressed, and what is the role of the affective dimension of infant-directed speech (here, fatherese)?

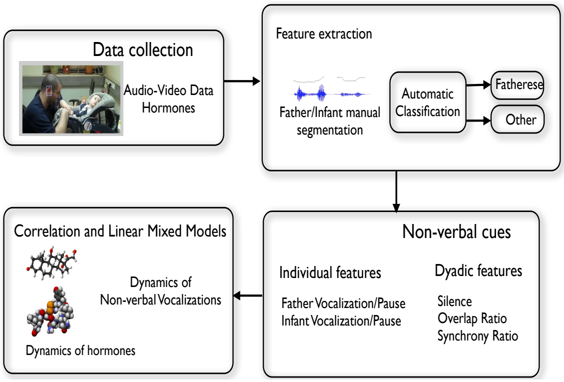

Fig. 1. The computational framework for the study of correlations between individual/dyadic non-verbal features and hormonal modulation

In addition, we tested whether intranasal OT administered to the father may modulate the father-infant STT characteristics and/or the father's vocalization (including fatherese) during interaction using a SF paradigm. To this end, a double-blind, placebo-controlled, crossover experimental design was employed. CT and OT were monitored during the experiment to assess the dynamics of these hormones (see Figure 2).

Given the background described above, we formulated the following hypotheses: (1) Before SF, we expected STT synchrony features, father vocalization and fatherese to increase under OT conditions and to be associated with OT elevation in infants. (2) Because the SF paradigm elicits a brief period of social stress, we also hypothesized that infants' vocalizations and CT will increase after SF. However, these increases, which represent a natural response to stress (e.g., an infant attempting to attract and maintain his fathers attention), could be modulated under OT conditions.

We also hypothesize that (3) after SF, fathers should regulate the interaction and adapt themselves to their child to re-establish the previous level of vocal synchrony. More generally, the current work sought to investigate the interplay between various non-verbal cues and hormonal changes.

4 DATA AND ANNOTATIONS

The current work exploits a protocol that was designed for investigating the biological and developmental dimensions of parent-infant interactions [39], [40]. Consequently, the interaction duration satisfied the requirements of both the SF paradigm and the timing needed to measure the hormones (see Figure 2).

4.1 Participants

Thirty-five healthy fathers (average age 29.7 years, s.d. = 4.2, range 22-38) participated with their five-month-old infants (s.d. = 1.25 months, range 4-8) in two laboratory visits, one week apart (total N = 70). Females were not enrolled in this study because of the physiological effects of OT manipulation (e.g., uterus contraction) and the need to control for menstrual cycle.

4.2 Procedure and OT administration

Following their arrival at the laboratory, fathers were asked to self-administer IU of either OT (Syntocinon Spray, Novartis, Basal, Switzerland) or a placebo. The administration order was counterbalanced, and participants and experimenters were blind to the conditions (OT or placebo administration). Forty minutes after administration, the infant joined his or her father in the observation room. The infant was seated in an infant seat mounted on a table. Father-infant interaction began approximately 45 min after substance administration.

4.3 Recording father-infant interaction

Each father-infant interaction lasted for 8 min: 3 min of free play, 2 min of parental SF and another 3 min of free play. Interactions were videotaped using a Flip Mino HD digital camcorder (Cisco, Irvine, CA) for offline coding of the behavior.

4.4 Salivary oxytocin and cortisol collection

A full description of salivary OT collection and ELISA analysis is provided in a previous publication by Weissman el al. [41]. Saliva samples from the fathers and infants were collected by Sallivatte (Sarstedt, Rommelsdorft, Germany) at multiple time-points (Figure 2): T1 (before inhaling) - before substance administration (father only); T2 (baseline) - prior to interaction, 40 min after OT administration; T3 (reactivity) - 20 min after interaction began; and T4 (recovery) - 20 min thereafter. The ratio of the change in the infant's salivary OT after and before the father-infant interaction (i.e., infant OT at T3/ infant OT at T2) was calculated. Log transformation was applied to this ratio, and the variable exhibited a normal distribution according to the Kolmogorov-Smirnov test.

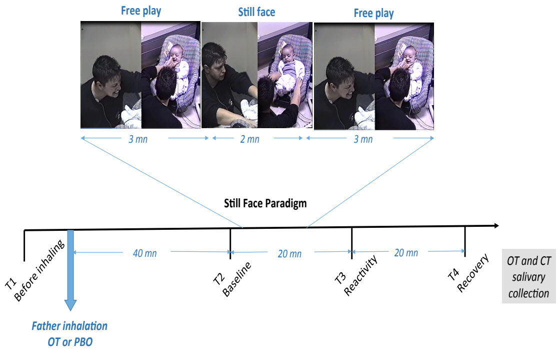

Fig. 2. Still-Face paradigm: dynamics of OT and CT salivary collection

CT levels were assayed using a commercial ELISA kit (Assay Design, MI, USA). Measurements were performed in duplicate according to the manufacturers instructions at multiple time-points, as done for OT. CT levels were calculated using Mathworks Inc Matlab (version 7.0) according to the relevant standard curves. CT change was defined as [T3 - T2].

5 NON-VERBAL FEATURE EXTRACTION

We present in this section the extracted non-verbal features that we hypothesized to be linked with the hormonal changes (Figure 1).

5.1 Speech turn-taking (STT)

As previously mentioned, the data employed here were collected in uncontrolled settings, making the use of voice activity detection systems difficult. Therefore, infants' and fathers' utterances were labeled according to the following categories by two annotators (blind to the conditions): *father vocalization* (meaningful vocalizations, laughing, singing, and animal sounds), *infant vocalization* (babbling vocalizations, laughing, and crying), and other *non-speech sounds* (clap the hand, snap fingers or the tongue, and mouth noise, for example). Cohen's kappa between the two annotators was calculated for each dyad, each task and each item of the grid. For all items, the kappa values were between 0.82 and 1. We obtained 20193 segments of vocalization (9706 before SF and 10487 after SF). The duration of the segments ranged from 0.1 and 7.79 s.

Based on this manual segmentation of the speakers' turns, we extracted several features to describe the dynamics of speech turns during the task:

- *Vocalization durations*. A speech turn is a continuous time segment during which one participant is speaking: *Father Vocalization* and *Infant Vocalization*. To describe the dynamics of these events, we computed standard statistics on the duration of the segments (mean, median, standard deviation, range, minimum and maximum).

- *Empty Pause Durations*. Silence >150 ms between two vocalizations. We extracted all pauses for each participant and calculated standard statistics on the duration of the segments.
- *Overlap Ratio*. We measured the percentage of interaction time when both father and infant were talking simultaneously.
- *Silence*. Silence is defined as sequences of time during which neither of the participants spoke for more than 150 ms. Statistics on silence events will provide insights regarding the dynamics of social interactions.
- *Synchrony Ratio*. This parameter is the number of infant's responses to his father vocalization within a time limit of 3 s divided by the number of father vocalizations during the time paradigm. The 3-s window was based on the available literature on synchrony [16], [42], [43].

By definition, during SF, fathers remain still, and thus, the use of STT is inappropriate. However, the infant is free to vocalize, so we applied the same annotation for each infant during SF and extracted his/her vocalizations.

5.2 Automatic fatherese detection

Previous works on parent-infant interaction have shown that parents vocalizations play a role in several developmental dimensions, such as bonding and language acquisition [44], [7]. These vocalizations, often termed infant-directed speech, are considered in their affective dimension nature and referred to as "parentese" ("fatherese" or "motherese") [7].

In this paper, we automatically detect those vocalizations (see Figure 1). Several statistical classifiers were designed [45]. The first one exploits a Gaussian mixture model (GMM), while the second one is a k-nearest neighbors (k-NN) classifier. For feature extraction, we employed Mel frequency cepstral features (MFCC), and the classifier was considered to be segmental due to the window length analysis employed for computing the features (20 ms). In contrast, when we used a set of functionals (mainly statistics: mean, maximum, minimum, variance and range) applied to the fundamental frequency, energy and duration of events, the classifier was considered to be a supra-segmental classifier because features are computed for only voiced segments. Each one of these classifiers is designed to produce posteriori probabilities: P_{seg} and P_{supra} [46], [47]. The classifiers were fused by a simple weighted sum of likelihoods for the different classifiers. That is, for a given utterance U_x, we compute:

$$C = \arg\max_{l}[\lambda \log P_{seg}(C_l|U_x)+(1-\lambda)\log P_{supra}(C_l|U_x)] \quad (1)$$

where $l = 1$ ("fatherese") or 2 ("other vocalization"). λ denotes the weighting coefficient.

To train the classifiers, we manually annotated 200 sequences from fathers (100 parentese vs. 100 other speech) that were validated by two psycholinguists [Cohen's kappa = 0.82; 95% Confidence Interval (CI) = 0.75-0.90]. The fathers were all speaking Hebrew. The classifiers were trained and evaluated using a 10-fold cross-validation scheme, and we compute:

- Accuracy, ACC =TruePositive/(Number of Examples)
- Positive Predictive Value, PPV =TruePositive/(TruePositive + FalsePositive)
- Negative Predictive Value, NPV (=FalsePositive/(TruePositive + FalsePositive)

PPV evaluates the proportion of correctly detected fatherese utterances out of all utterances labelled as fatherese. PPV can be viewed as the reliability of positive predictions induced by the classifier. NPV gives similar information regarding the detection of other vocalizations.

The system's performance was as follows: accuracy= 84% (95%CI= 64.27-92.26%); PPV= 79.3% (95%CI= 57.79-82.70%); and NPV= 90.47% (95%CI= 62.16- 88.53%). In its most effective configuration, the detector used only the GMM classifier for both segmental and supra-segmental features (M, number of Gaussians for the GMM Classifier: M=12 and M=15, respectively, and λ=weighting coefficient used in the equation fusion: λ=0.4). We also evaluated the system in a speaker-dependent condition: accuracy= 87.5% (95%CI= 82.91-92.08%); PPV= 88.47% (95%CI= 83.03-95.18%); and NPV= 86.41% (95%CI= 79.4-92.88%) [45].

Based on the automatic detection of fatherese, we extracted several features to describe the dynamics of speech turns:

- *Fatherese Ratio*. Duration of fatherese vocalization/Duration of all father's vocalizations.
- *Non Fatherese Ratio*. Duration of non fatherese vocalization/ Duration of all father's vocalizations.
- *Synchrony Fatherese Ratio*. Ratio of time during which the infant vocalizes in response to his/her father fatherese.
- *Synchrony Non Fatherese Ratio*. Ratio of time during which the infant vocalizes in response to his/her father other vocalizations.

As done for speech turns, we extracted statistical features: mean, median, standard deviation, range, minimum and maximum.

6 CORRELATION ANALYSIS

6.1 Methods

We performed a linear mixed effects analysis of the relationship between hormonal and non-verbal vocalization dynamics by considering the following conditions: group OT vs. placebo, visit 1 vs. visit 2, and before vs. after SF. The LMMs were fit using R Core Team [48].

The statistical Shapiro-Wilk test was applied in order to determine whether the variables follow a normal distribution. In the case of the Synchrony Ratio, because the variables did not show normal distributions, a transformation was necessary before processing the LMM (see 6.2). The Overlap Ratio variable was assessed by a paired Wilcoxon nonparametric test because of both the non-normal distribution and the number of samples. Predictions of infant OT increase according to the father's turn-taking variables were also performed with LMM.

The LMM was expressed as:

$$y_{ij} = \mu + \beta_1 I_{ij} + \beta_2 O_{ij} + \beta_3 T_{ij} + \gamma D_i + \epsilon_{ij} \qquad (2)$$

where y_{ij} is the outcome variable, which is a nonverbal feature (vocalization, pause, synchrony parameter, cf. section 5) computed for the dyad D_i for each visit $j = 1, 2$. The other predictor variables are the treatment condition I_{ij} (OT or placebo group), the time T_{ij} (before, during and after SF, see section 4.4) for each dyad D_i and the administration order O_{ij}, which was counterbalanced. Participants and experimenters were blind to the conditions (OT or placebo administration and visits).

LMMs allow estimations of the effect of a variable considering the other variables fixed or not. Thus, we fixed the effect of each independent variable (treatment, visit, and SF) and estimated their effect of each on each dependent variable (vocalization, pause, and synchrony parameters).

To assess infant vocalization change during the three study periods (before SF, during SF and after SF), a secondary analysis using the same LMM was used with three independent variables: group I_{ij} (OT vs. PL), order (visit $j = 1, 2$) and time T_{ij} (before, during and after SF). In the case of significant results, post hoc analyses were conducted to compare infant vocalization change in terms of consecutive temporal segments.

OT and CT variations were also explored using LMM to assess the effect of treatment group and/or time (see fig 2). To assess how the CT and OT baseline values were associated with STT variables and fatherese, we used correlation analyses within OT and CT baseline values and STT variables during the first free play period before SF under the placebo condition. Finally, we explored the association between the dynamics of OT and CT (e.g., baseline values and changes) and STT variables with time T_{ij} and treatment condition I_{ij} in a LMM.

6.2 Results

6.2.1 Description of the data

We computed the features described in section 5 to describe (1) individual dynamics of fathers (*Vocalization*, *Pauses*, *Fatherese Ratio* and *Non-Fatherese Ratio*) and children (*Vocalization* and *Empty Pauses*); and (2) dyadic dynamics, such as STT: *Silence*, according to the *Synchrony Ratio*, *Synchrony Fatherese Ratio*, *Synchrony Non-Fatherese Ratio*, and *Overlap Ratio*. For each feature, we computed standard statistics (mean, median, standard deviation, range, minimum and maximum), which are employed as outcome variables of LMM.

Table 2 presents the descriptive statistics of the speech variables we selected. We list them separately under OT or placebo condition and before or after SF.

Figure 3 summarizes the dynamics of OT and CT during the experiment. As previously reported [39], there was a significant treatment and time effect for OT in both fathers (who inhaled OT: time effect: $p < 0.001$; treatment effect: $p < 0.001$; time x treatment effect: $p < 0.001$) and infants (time effect: $p < 0.001$; treatment effect: $p < 0.001$; time x treatment effect: $p < 0.001$). Father CT showed a significant time effect ($p < 0.001$) with a higher value at baseline and before inhaling than during reactivity or recovery but no treatment effect ($p = 0.97$). Contrary to our expectations, infant CT did not show treatment ($p = 0.09$) or time effects ($p = 0.15$).

6.2.2 Variation of STT and fatherese according to treatment and time

Because of the possible different influence of the three variables (group OT vs. placebo, visit 1 vs. visit 2, and before vs. after SF), we used LMM to assess their respective associations with father/infant STT and fatherese variables. Thus, we fixed the effect of the each independent variable (treatment, visit, SF) and estimated their own effect on each dependent variable (vocalization, pause, synchrony parameters). The results are presented in Table 3.

Despite our hypotheses, we found very few significant associations between OT and STT parameters in fathers or infants or between OT and STT synchrony. Only the maximum duration of *Father Empty Pause* was significantly higher and the *Non-Fatherese Ratio* was significantly lower under OT. However, the mean durations of both *Father Pause* and *Father Vocalization* were not significantly different whether fathers received OT or placebo. The maximum duration of *Silence* was significantly lower at the second visit, but we found no effect on mean duration. In addition, *Fatherese Ratio* was significantly higher at the second visit.

We concluded that the effect of the visit had a limited influence on STT variables during father-infant interaction, although fathers anticipated the

TABLE 2
Description of the speech turn taking and fatherese variables during father/infant interaction before and after still face, and under Oxytocin or Placebo conditions

Before Still Face		
Father parameters	Oxytocin	Placebo
Vocalization: Mean (SD) Max	1.37 (0.61) 3.27	1.38 (0.63) 2.75
Empty Pause: Mean (SD) Max	1.61 (0.87) 4.92	1.45 (0.67) 3.86
Fatherese ratio: Mean (SD) Max	0.06 (0.09) 0.44	0.05 (0.08) 0.37
Non Fatherese ratio: Mean (SD) Max	0.39 (0.16) 0.68	0.42 (0.17) 0.71
Infant parameters		
Vocalization: Mean (SD) Max	0.74 (0.31) 1.28	0.75 (0.55) 2.87
Empty Pause: Mean (SD) Max	1.01 (1.02) 4.54	1 (0.84) 3.91
Speech turn taking synchrony		
Silence: Mean (SD) Max	1.70 (0.9) 4.49	1.54 (0.79) 3.86
Overlap Ratio: Mean (SD) Max	0.03 (0.05) 0.21	0.03 (0.06) 0.29
Synchrony Ratio: Mean (SD) Max	0.28 (0.19) 0.64	0.3 (0.19) 0.77
Synchrony Ratio > Fatherese: Mean (SD) Max	0.32 (0.3) 1	0.43 (0.36) 1
Synchrony Ratio > Non Fatherese: Mean (SD) Max	0.27 (0.19) 0.64	0.29 (0.19) 0.66
After Still Face		
Father parameters	Oxytocin	Placebo
Vocalization: Mean (SD) Max	1.24 (0.39) 2.08	1.29 (0.48) 2.65
Empty Pause: Mean (SD) Max	1.45 (0.81) 4.04	1.32 (0.72) 3.54
Fatherese ratio: Mean (SD) Max	0.06 (0.08) 0.41	0.05 (0.07) 0.29
Non Fatherese ratio: Mean (SD) Max	0.37 (0.15) 0.66	0.41 (0.17) 0.65
Infant parameters		
Vocalization: Mean (SD) Max	1.42 (1.32) 5.80	1.56 (1.56) 7.15
Empty Pause: Mean (SD) Max	1.19 (1.27) 6.41	0.92 (0.63) 2.47
Speech turn taking synchrony		
Silence: Mean (SD) Max	1.47 (0.73) 3.59	1.34 (0.86) 4.46
Overlap Ratio: Mean (SD) Max	0.07 (0.09) 0.3	0.09 (0.14) 0.6
Synchrony Ratio: Mean (SD) Max	0.4 (0.25) 1	0.41 (0.25) 1
Synchrony Ratio > Fatherese: Mean (SD) Max	0.44 (0.33) 1	0.55 (0.36) 1
Synchrony Ratio > Non Fatherese: Mean (SD) Max	0.39 (0.25) 1	0.41 (0.26) 1

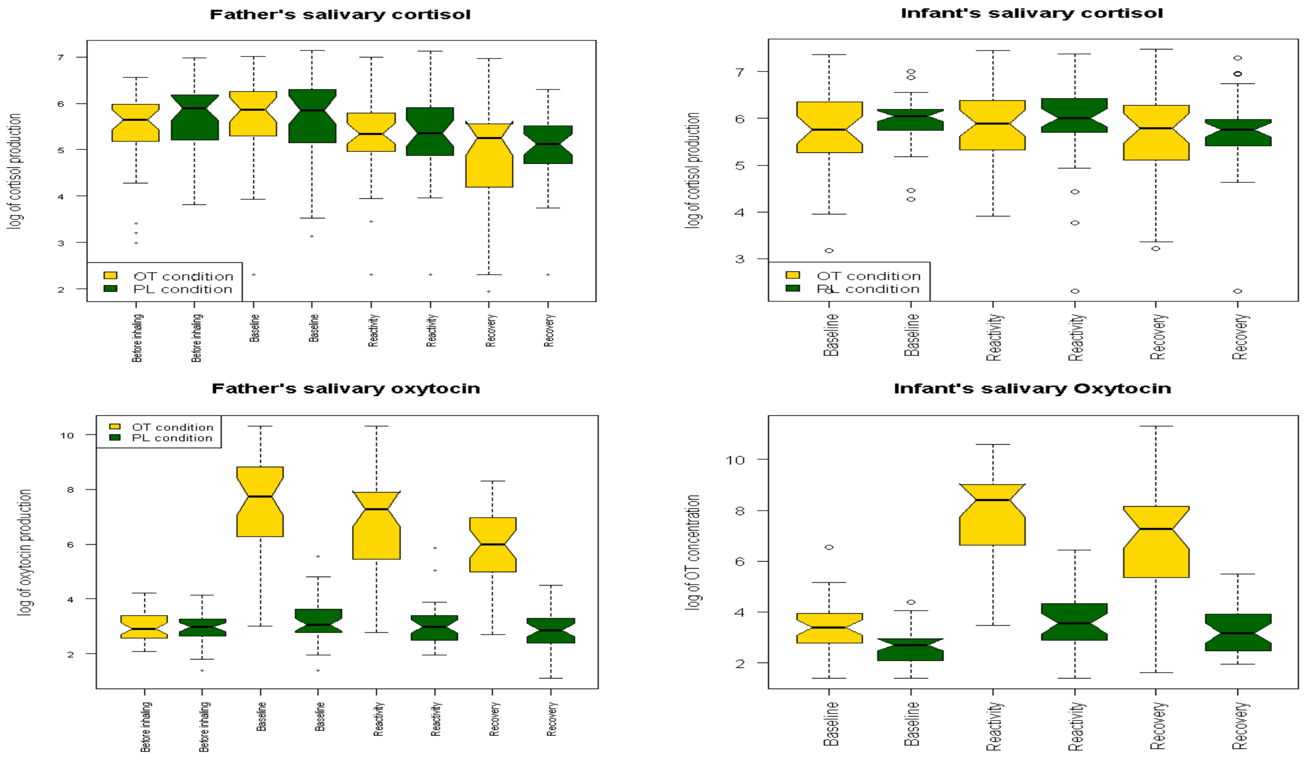

Fig. 3. Father (left) and infant (right) hormonal changes during the experiment: In this figure, we present the course of father (left) and infant (right) hormonal values during the experiment (cortisol is up, oxytocin is down). Values under placebo condition are given in green, where values under oxytocin condition are given in yellow.

TABLE 3
Linear Mixed Model of speech turn taking and fatherese during father infant interaction using a still face paradigm under Oxytocin or Placebo

	Father parameters		
	Oxytocin effect	Visit effect	Still face effect
Father Vocalization Mean	-0.01 (p=0.79)	0.01 (p=0.79)	-0.05 (p=0.17)
Father Vocalization Max	-0.075 (p=0.26)	-0.01 (p=0.86)	-0.11 (p=0.11)
Father Pause Mean	0.08 (p=0.24)	-0.05 (p=0.46)	**-0.14 (p=0.027)**
Father Pause Max	**0.24 (p=0.046)**	-0.09 (p=0.53)	-0.21 (p=0.078)
Fatherese Ratio	0.22 (p=0.83)	**2.67 (p=0.0143)**	0.38 (p=0.72)
Non Fatherese Ratio	**-0.03 (p=0.044)**	-0.02 (p=0.17)	-0.015 (p=0.32)
	Infant parameters		
	Oxytocin effect	Visit effect	Still face effect
Infant Vocalization Mean	-0.05 (p=0.59)	-0.06 (p=0.52)	**0.46 (p<0.001)**
Infant Vocalization Max	-0.039 (p=0.76)	-0.01 (p=0.9)	**0.69 (p<0.001)**
Infant Pause Mean	0.05 (p=0.69)	0.07 (p=0.62)	-0.006 (p=0.96)
Infant Pause Max	0.17 (p=0.43)	0.11 (p=0.61)	0.08 (p=0.69)
	Speech Turn Taking and Synchrony		
Silence Ratio Mean	0.1 (p=0.086)	-0.11 (p=0.067)	-0.17 (p=0.007)
Silence Ratio Max	0.26 (p=0.065)	-0.3 (p=0.033)	-0.31 (p=0.027)
Overlap Ratio Mean	V=213.5 (p=0.82)	Invalid	**V=224 (p<0.001)**
Synchrony Ratio	-0.01 (p=0.68)	-0.02 (p=0.40)	**0.09 (p<0.001)**
Synchrony Fatherese Ratio	-0.11 (p=0.12)	-0.06 (p=0.36)	0.12 (p=0.11)
Synchrony Non Fatherese Ratio	-0.02 (p=0.63)	-0.04 (p=0.28)	**0.13 (p<0.001)**

experiment at visit 2 and used less silence and more fatherese.

This was not the case for the SF variable (last column Table 3). Infant mean vocalizations were longer after SF than before SF ($\beta = 0.46, p < 0.001$). The maximum duration of infant vocalization was also significant ($\beta = 0.69, p < 0.001$), while the infant pause characteristics did not change. For the father, we did not observe any significant changes in vocalization (either mean or maximum). However, we did find a significant decrease in father pause characteristics after SF. The mean duration of *Empty Pauses* was shorter after SF than before SF ($\beta = -0.14, p = 0.027$), and there was a trend in the maximum duration ($\beta = -0.21, p = 0.078$).

Regarding STT synchrony parameters, we found that the mean ($\beta = -0.17, p = 0.007$) and maximum ($\beta = -0.31, p = 0.027$) durations of *Silence* were shorter after SF than before SF. Because the visit effect was significant, we introduced both visit (visit 1 vs. visit 2) and time (before SF vs. after SF) into the LMM. A significant difference was found ($\beta = 0.8, p = 0.003$), suggesting a larger difference in in the silence maximum value between before and after SF during the second visit.

Given the non-normal distribution of the *Synchrony Ratio*, the following transformation was performed: $\sqrt{(SynchronyRatio + 0.1)}$. The association with the SF paradigm was significant ($\beta = 0.09, p < 0.001$), meaning that there was an increase in the *Synchrony Ratio* after SF. Further analysis showed that the increase was particularly evident for the *Synchrony Non-Fatherese Ratio*, meaning that the infants response to the father's vocalizations was higher after SF, even when fathers did not use fatherese. Finally, the *Overlap Ratio* was significantly increased after SF (as assessed by a paired Wilcoxon non-parametric test).

To explore whether the increase in the mean duration of infant vocalization occurred during SF, the LMM was employed before, during and after SF. There was no effect of OT treatment ($\beta = 0.005, p = 0.95$) or order of visit ($\beta = 0.037, p = 0.64$). There was a significant effect of time (during SF: $\beta = 0.41, p < 0.001$; after SF: $\beta = 0.46, p < 0.001$). Post hoc analyses showed that an increase in infant vocalization occurred during SF, and a significant difference was observed between the during-SF and before-SF values ($\beta = 0.415, p < 0.001$) but not between the during-SF and after-SF values ($\beta = -0.044, p = 0.89$). However, the difference remained significant between the before-SF and after-SF values ($\beta = 0.46, p < 0.001$).

6.2.3 Association of OT and CT baseline values and changes with STT variables

To assess whether STT variables predicted infant OT increases after father-infant interaction, we used LLM with the log of OT increase as the variable to be explained. No significant association other than whether the father was exposed to OT was found with the STT variables.

We investigated correlations between OT and CT baseline values and STT variables during the first period of free play before SF under the placebo condition only. Indeed, massive OT inhalation in fathers and the resulting infant OT increase prevent any interpretation of correlations with previous OT or CT values.

Under placebo conditions, father salivary CT at baseline was correlated with better *Synchrony Ratio* ($\rho = 0.43, p = 0.015$) and with the *Fatherese Ratio*

($\rho = 0.42, p = 0.013$) and *Non-Fatherese Synchrony* ($\rho = 0.42, p = 0.018$). We explored STT variables showing a significant association with time and/or condition using LMM, including OT and CT baseline values and changes in the model. We found a significant association for infant vocalization with time. Infant vocalization showed an increase after SF ($\beta = 0.46, p < 0.001$) and no effect of treatment condition. Additionally, no effect of hormones was observed, except a tendency to be positively associated with Δ infant CT ($\beta = 0.00026, p = 0.065$).

7 DISCUSSION

7.1 OT does not shape vocalization, including fatherese

In contrast to our hypothesis, OT did not influence STT, father vocalization, or synchrony parameters during the free play or during the reunion after the SF social stressor. OT appeared to slightly increase the fathers maximal duration of empty pauses, but this single effect is too marginal to allow further interpretation. Based on previous works showing differential response to the SF stressor in infants' CT response [49] and infants' vagal regulation [36], we also estimated the prediction of infants' increased OT according to fathers' STT variables. None of the fathers' parameters were significantly linked with infants' OT levels.

Previous works have shown that (1) parentese plays a major role in early infant-caregiver interaction [7] and (2) OT production and action are both modulated by the emotional valence of the stimuli [50]. We investigated the effect of fatherese vocalization and did not find any effect of OT. However, fathers seemed to anticipate the experiment at the second visit (see section 4) and used less silence and more fatherese. These findings suggest that the dynamics of vocalization features, including fatherese, are not correlated with OT, although they are both essential in social interaction and infant-caregiver relation. Previously, we have shown that OT administration increased affectionate touch [39] and shaped parental motion during father-infant interaction [40]. Building on these findings, the current study suggests that OT modulates physical and visual cues rather than speech/audio cues.

To understand these results that challenge our hypothesis, we propose the following interpretations. First, the influence of OT on parenting emotion and behavior is modulated by several contextual and inter- and intra-individual parameters. This is also the case for infant response [49].

Second, it is likely that changes in social cues for bonding during evolution (see also Table 1) [51] that are shaped by OT may occur during development. Indeed, the study of Seltzer et al. [38], which showed that social vocalizations could release OT in humans, was performed in much older subjects (children aged 7-12 years). Similarly, in a previous work [40], we did not find that the distance between father and infant during free play interaction was shaped by OT but that motion and acceleration were [40], while Scheele et al. [52] reported that OT modulates social distance between adult males and females, promoting fidelity within monogamous human relationships. These examples illustrate that, despite the key role of OT neuroendocrine regulation of social interaction and partnership, the social cues involved may differ according to age.

Regarding CT, we expected a significant time interaction in infants. This was not the case, although the standard deviation values were larger (Figure 3). In contrast, we observed a significant time interaction in fathers, meaning that fathers were stressed by participating in the experiment, which resulted in some type of anticipation, whether under OT or placebo. This finding may explain why father salivary CT at baseline was correlated with a higher fatherese ratio and better synchrony and non-fatherese synchrony, which means that fathers used more fatherese during free play.

7.2 Vocalization and STT across the SF paradigm

We summarized in Figure 4 the role of infant vocalization, father vocalization and STT during the SF paradigm, taking into account the significant results and time course. First, fathers remained still, and infant vocalizations (mean and maximum durations) increased in response to this dialogue interruption, which occurred after the first 3 min of free play. The increase in infant mean and maximum empty pause durations can be related to the silence of the fathers during this period.

This result is in line with those of previous studies. Based on several clinical observations, infant behavior modification during SF includes emotional changes (decreased smiling and increased signals of distress), behavior encouraging the caregiver to re-engage in interaction (vocalizing and gesturing at the caregiver), and averted gaze when efforts to re-engage in interaction fail [53], [54]. Two studies focusing on infant vocalization confirmed that during SF, infants' vocalizations were a way to revive infant/caregiver dialog [55], [53]. They were able to identify two functions for this signal: emotional expression/reaction and interactional regulation. Here, using SSP methods, we delve even deeper and describe infant-father adaptation during reunion after SF in terms of both each partners vocalization and STT variables.

After SF, infant mean vocalization duration remained increased, but the maximum duration was shorter. The mean and maximum duration of Empty Pauses were also shorter. We can conclude that infants continue to produce vocalizations after SF. We argue that the effect of SF on infants' vocalizations during the reunion episode may be explained by the persistence of the distress. Although not significant, the statistical tendency associating increased infant vocalization and CT changes further supports this argument. This finding helps to confirm that infants use vocalization as a way to regulate and restore interaction.

Fathers' empty pauses and vocalization characteristics reflected the dynamics of paternal adaptation. After SF, the mean duration of empty pauses was shorter, but the mean, maximum durations of vocalization and the maximum of empty pauses did not increase. Fathers seemed to produce more frequent vocalizations but shorter empty pauses after SF than before. This adaptation may provide more opportunities for infants to interact with them (more empty pauses) while maintaining a certain interaction dynamic (shorter duration of vocalization). Fathers seemed to intend to elicit infant responses without overstimulation. They contributed to supporting and helping infants in re-establishing synchrony by modulating the dynamics of their own vocalization. Sravish et al. [35] also reported an increase in dyadic flexibility after an SF episode. Those findings highlight the regulatory function of dyads, which tend to spontaneously repair the broken synchrony.

The analysis of STT synchrony parameters, combined with the previous analysis, allowed us to highlight the mutual contingency between partners and their reciprocal non-verbal regulation. Indeed, the mean value of the *Silence ratio* was lower after SF, showing more mutual responses between both partners. The increase in the infants' responses to their fathers was confirmed by the synchrony ratio, which was highly significant here. Knowing that infants' vocalizations are more frequent after SF and that fathers make more frequent but briefer empty pauses, the *Overlap Ratio* describes the difficulty experienced by partners in mutually adapting their interactions during the reunion time.

7.3 Limitations, strengths and research to be done

The current study should be interpreted within the context of its limitations and strengths. The limitations include (i) the sample size; (ii) the experimental nature of the interaction and the cross-over design that induced a visit effect; (iii) the bias regarding volunteer and motivated fathers, meaning that generalization may be speculative; and (iv) differences in time points regarding SF course and hormonal dosage because we chose saliva samples.

The strengths include (i) the choice of fathers for the experiment, who do not experience lactation-related OT variation as mothers do, (ii) the homogeneous age of the infants, (iii) the experimental design allowing LMM and blind segmentation, (iv) the use of SSP to assess both low and high level audio features, and (v) the multimodal approach that included physiological signals (e.g., hormones).

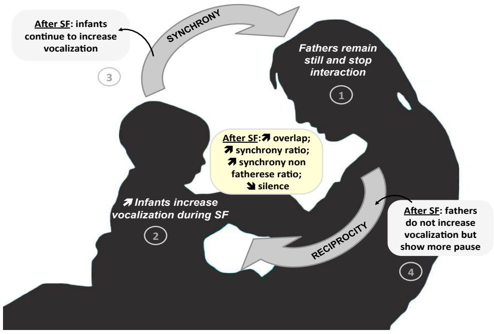

Fig. 4. Impact of father SF on interaction dynamics in terms of speech turn-taking. In this figure, we summarize the effect of father SF on the interaction dynamics in terms of speech turn-taking. (1) Fathers remain still and stop interaction. (2) During the SF, infants increase their vocalizations. (3) When interaction restarts after SF, infants maintain a high level of vocalization, which impacts the synchrony parameters. (4) After SF, silence decreases and the overlap and synchrony ratios increase, particularly the synchrony non-fatherese ratio.

8 CONCLUSION AND FUTURE WORK

In this work, we proposed a computational framework for the evaluation of parent-infant interaction using non-verbal cues and hormonal modulation. To the best of our knowledge, this is the first attempt to automatically detect and analyze the interplay between hormonal dynamics and non-verbal behaviors.

The long-term goal is to develop a new framework that exploits cues that have not been previously considered in automatic behavior analysis. Our research focus is two-fold: 1) to systematically examine social bonding and synchrony between the two partners, and 2) to automatically analyze the interplay between hormonal modulation dynamics and both speech turns and affective speech ("fatherese").

In the future, we plan to extend modeling of the interplay between the various multimodal cues and hormonal modulation. This can be achieved by investigating other paradigms, such as those of adults, that are usually employed by the automatic behavior analysis research community. From a theoretical perspective, it will be interesting to compare modeling-based results to traditional psychometric measurements (e.g., manual annotation and questionnaires) and hormonal changes to investigate interaction quality. We are currently evaluating multi-task learning schemes for this purpose.

Recently, Meltzoff et al. [56] described how research in developmental psychology may provide a good opportunity for enhancing computational models of such phenomena and vice-versa. We conclude that the methods developed for this study may be applied in other contexts and that is also opens new avenues of research applications.

ACKNOWLEDGMENTS

The study was supported by the German-Israeli Foundation grant (no. 1114-101.4/2010) to R.F., by the Agence Nationale de la Recherche (ANR-12-SAMA-006) to D.C., and the Groupement de Recherche en Psychiatrie (GDR-3557) to D.C. This work was partially performed within the Labex SMART (ANR-11-LABX-65) supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02. Sponsors had no involvement in study design, data analysis or interpretation of results. The research was approved by the Institutional Review Board at Bar-Ilan University, was conducted according to ethical standards, and all participating fathers signed an informed consent.

REFERENCES

[1] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[2] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition (FG'11)*, Santa Barbara, CA, USA, March 2011.

[3] A. Tarasov, C. Cullen, and S. Delany, "Using crowdsourcing for labeling emotional speech assets," in *W3C workshop on Emotion ML.*, 2010.

[4] M. Soleymani and M. Pantic, "Human-centered implicit tagging: Overview and perspectives," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, Oct 2012, pp. 3304–3309.

[5] M. Nagasawa, S. Okabe, M. K., and T. Kikusui, "Oxytocin and mutual communication in mother-infant bonding," *Frontiers in Human Neuroscience*, vol. 6, 2012.

[6] B. Thierry, "Identifying constraints in the evolution of primate societies," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, no. 1618, 2013.

[7] C. Saint-georges, M. Chetouani, R. Cassel, F. Apicella, A. Mahdhaoui, F. Muratori, M. Laznik, and D. Cohen, "Motherese in interaction: at the cross-road of emotion and cognition? (a systematic review)," *PLoS ONE*, vol. 8, no. 10, p. e78103, 2013.

[8] I. Branchi, D. Santucci, and E. Alleva, "Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development," *Behav. Brain Res.*, vol. 125, no. 1-2, pp. 49–56, Nov 2001.

[9] D. Falk, "Prelinguistic evolution in early hominins: whence motherese?" *Behav Brain Sci*, vol. 27, no. 4, pp. 491–503, Aug 2004.

[10] F. Lévy, M. Keller, and P. Poindron, "Olfactory regulation of maternal behavior in mammals," *Hormones and Behavior*, vol. 46, no. 3, pp. 284–302, 9 2004.

[11] D. Bjorklund, "A note on neonatal imitation," *Developmental Review*, vol. 7, no. 1, pp. 86–92, 3 1987.

[12] R. Franceschini, P. L. Venturini, A. Cataldi, T. Barreca, N. Ragni, and E. Rolandi, "Plasma beta-endorphin concentrations during suckling in lactating women." *Br J Obstet Gynaecol*, vol. 96, no. 6, pp. 711–713, Jun 1989.

[13] S. Moriceau and R. M. Sullivan, "Unique neural circuitry for neonatal olfactory learning." *J Neurosci*, vol. 24, no. 5, pp. 1182–1189, Feb 2004.

[14] K. Lorenz, "Der kumpan in der umwelt des vogels," *Journal für Ornithologie*, vol. 83, no. 2, pp. 137–213, 1935.

[15] A. Paukner, P. F. Ferrari, and S. J. Suomi, "Delayed imitation of lipsmacking gestures by infant rhesus macaques (macaca mulatta)." *PLoS One*, vol. 6, no. 12, p. e28848, 2011.

[16] R. Feldman, "Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions." *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 48, no. 3-4, pp. 329–354, 2007.

[17] R. Feldman, A. Weller, L. Sirota, and A. I. Eidelman, "Testing a family intervention hypothesis: the contribution of mother-infant skin-to-skin contact (kangaroo care) to family interaction, proximity, and touch." *J Fam Psychol*, vol. 17, no. 1, pp. 94–107, Mar 2003.

[18] J. K. Rilling, "The neural and hormonal bases of human parental care." *Neuropsychologia*, vol. 51, no. 4, pp. 731–747, Mar 2013.

[19] K. Macdonald and T. M. Macdonald, "The peptide that binds: a systematic review of oxytocin and its prosocial effects in humans," *Harv Rev Psychiatry*, vol. 18, no. 1, pp. 1–21, 2010.

[20] R. Feldman, I. Gordon, and O. Zagoory-Sharon, "Maternal and paternal plasma, salivary, and urinary oxytocin and parent–infant synchrony: considering stress and affiliation components of human bonding," *Developmental Science*, vol. 14, no. 4, pp. 752–761, 2011.

[21] C. S. Carter, "Oxytocin pathways and the evolution of human behavior," *Annual Review of Psychology*, vol. 65, no. 1, pp. 17–39, 2014.

[22] O. Weisman and R. Feldman, "Oxytocin effects on the human brain: Findings, questions, and future directions," *Biological Psychiatry*, vol. 74, no. 3, pp. 158–159, 2013.

[23] A. Campbell, "Oxytocin and human social behavior," *Pers Soc Psychol Rev*, vol. 14, no. 3, pp. 281–295, Aug 2010.

[24] A. Levine, O. Zagoory-Sharon, R. Feldman, and A. Weller, "Oxytocin during pregnancy and early postpartum: individual patterns and maternal-fetal attachment," *Peptides*, vol. 28, no. 6, pp. 1162–1169, Jun 2007.

[25] A. R. Tarullo and M. R. Gunnar, "Child maltreatment and the developing HPA axis," *Horm Behav*, vol. 50, no. 4, pp. 632–639, Nov 2006.

[26] C. E. Hostinar, R. M. Sullivan, and M. R. Gunnar, "Psychobiological mechanisms underlying the social buffering of the hypothalamic-pituitary-adrenocortical axis: a review of animal models and human studies across development," *Psychol Bull*, vol. 140, no. 1, pp. 256–282, Jan 2014.

[27] D. H. Hellhammer, S. Wust, and B. M. Kudielka, "Salivary cortisol as a biomarker in stress research," *Psychoneuroendocrinology*, vol. 34, no. 2, pp. 163–171, Feb 2009.

[28] K. Dedovic, A. Duchesne, J. Andrews, V. Engert, and J. C. Pruessner, "The brain and the stress axis: the neural correlates of cortisol regulation in response to stress," *Neuroimage*, vol. 47, no. 3, pp. 864–871, Sep 2009.

[29] R. Feldman, M. Singer, and O. Zagoory, "Touch attenuates infants' physiological reactivity to stress," *Dev Sci*, vol. 13, no. 2, pp. 271–278, Mar 2010.

[30] R. Feldman, A. Granat, C. Pariente, H. Kanety, J. Kuint, and E. Gilboa-Schechtman, "Maternal depression and anxiety across the postpartum year and infant social engagement, fear regulation, and stress reactivity," *J Am Acad Child Adolesc Psychiatry*, vol. 48, no. 9, pp. 919–927, Sep 2009.

[31] B. Ditzen, M. Schaer, B. Gabriel, G. Bodenmann, U. Ehlert, and M. Heinrichs, "Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict," *Biol. Psychiatry*, vol. 65, no. 9, pp. 728–731, May 2009.

[32] M. Kanat, M. Heinrichs, and G. Domes, "Oxytocin and the social brain: Neural mechanisms and perspectives in human research," *Brain Res.*, Nov 2013.

[33] I. Gordon, O. Zagoory-Sharon, J. F. Leckman, and R. Feldman, "Oxytocin, cortisol, and triadic family interactions," *Physiology & Behavior*, vol. 101, no. 5, pp. 679 – 684, 2010.

[34] E. Tronick, H. Als, L. Adamson, S. Wise, and T. Brazelton, "The infant's response to entrapment between contradictory messages in face-to-face interaction," *Journal of American Academy of Child Psychiatry*, pp. 1–13, 1978.

[35] A. V. Sravish, E. Tronick, T. Hollenstein, and M. Beeghly, "Dyadic flexibility during the face-to-face still-face paradigm:

a dynamic systems analysis of its temporal organization," *Infant Behav Dev*, vol. 36, no. 3, pp. 432–437, Jun 2013.
[36] G. A. Moore, J. F. Cohn, and S. B. Campbell, "Infant affective responses to mother's still face at 6 months differentially predict externalizing and internalizing behaviors at 18 months," *Dev Psychol*, vol. 37, no. 5, pp. 706–714, Sep 2001.
[37] G. A. Moore and S. D. Calkins, "Infants' vagal regulation in the still-face paradigm is related to dyadic coordination of mother-infant interaction," *Dev Psychol*, vol. 40, no. 6, pp. 1068–1080, Nov 2004.
[38] L. J. Seltzer, T. E. Ziegler, and S. D. Pollak, "Social vocalizations can release oxytocin in humans," *Proc. Biol. Sci.*, vol. 277, no. 1694, pp. 2661–2666, Sep 2010.
[39] O. Weisman, O. Zagoory-Sharon, and R. Feldman, "Oxytocin administration to parent enhances infant physiological and behavioral readiness for social engagement," *Biological Psychiatry*, vol. 72, no. 12, pp. 982 – 989, 2012.
[40] O. Weisman, E. Delaherche, M. Rondeau, M. Chetouani, D. Cohen, and R. Feldman, "Oxytocin shapes parental motion during father–infant interaction," *Biology Letters*, vol. 9, no. 6, 2013.
[41] O. Weisman, O. Zagoory-Sharon, and R. Feldman, "Intranasal oxytocin administration is reflected in human saliva," *Psychoneuroendocrinology*, vol. 37, no. 9, pp. 1582 – 1586, 2012.
[42] C. Saint-Georges, A. Mahdhaoui, M. Chetouani, R. S. Cassel, M.-C. Laznik, F. Apicella, P. Muratori, S. Maestro, F. Muratori, and D. Cohen, "Do parents recognize autistic deviant behavior long before diagnosis? taking into account interaction using computational methods," *PLoS ONE*, vol. 6, no. 7, p. e22393, 07 2011.
[43] E. Delaherche, M. Chetouani, M. Mahdhaoui, C. Saint-Georges, S. Viaux, and D. Cohen, "Interpersonal synchrony : A survey of evaluation methods across disciplines," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 349– 365, 2012.
[44] D. Cohen, R. S. Cassel, C. Saint-Georges, A. Mahdhaoui, M.-C. Laznik, F. Apicella, P. Muratori, S. Maestro, F. Muratori, and M. Chetouani, "Do parentese prosody and fathers' involvement in interacting facilitate social interaction in infants who later develop autism?" *PLoS ONE*, vol. 8, no. 5, p. e61402, 05 2013.
[45] A. Mahdhaoui, M. Chetouani, R. Cassel, C. Saint-Georges, E. Parlato, M. Laznik, F. Apicella, F. Muratori, S. Maestro, and D. Cohen, "Computerized home video detection for motherese may help to study impaired interaction between infants who become autistic and their parents," *International Journal of Methods in Psychiatric Research*, 2011.
[46] R. Duda, P. Hart, and D. Stork, *Pattern Classification, second edition.* Wiley, 2000.
[47] S. Kim, P. Georgiou, S. Lee, and S. Narayanan, "Real-time emotion detection system using speech: Multi-modal fusion of different timescale features," in *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on*, 2007, pp. 48–51.
[48] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013. [Online]. Available: http://www.R-project.org/
[49] O. Weisman, O. Zagoory-Sharon, and R. Feldman, "Oxytocin administration alters HPA reactivity in the context of parent-infant interaction," *Eur Neuropsychopharmacol*, vol. 23, no. 12, pp. 1724–1731, Dec 2013.
[50] M. M. Riem, M. H. van IJzendoorn, M. Tops, M. A. Boksem, S. A. Rombouts, and M. J. Bakermans-Kranenburg, "No laughing matter: intranasal oxytocin administration changes functional brain connectivity during exposure to infant laughter," *Neuropsychopharmacology*, vol. 37, no. 5, pp. 1257–1266, Apr 2012.
[51] K. D. Broad, J. P. Curley, and E. B. Keverne, "Mother-infant bonding and the evolution of mammalian social relationships," *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, vol. 361, no. 1476, pp. 2199–2214, Dec 2006.
[52] D. Scheele, N. Striepens, O. Gunturkun, S. Deutschlander, W. Maier, K. M. Kendrick, and R. Hurlemann, "Oxytocin modulates social distance between males and females," *J. Neurosci.*, vol. 32, no. 46, pp. 16 074–16 079, Nov 2012.
[53] C. E. Delgado, D. S. Messinger, and M. E. Yale, "Infant responses to direction of parental gaze: A comparison of two still-face conditions," *Infant Behavior and Development*, vol. 25, no. 3, pp. 311 – 318, 2002.
[54] Z. Hammal, J. F. Cohn, D. S. Messinger, W. I. Mattson, and M. H. Mahoor, "Head movement dynamics during normal and perturbed parent-infant interaction," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013, Geneva, Switzerland, September 2-5, 2013*, 2013, pp. 276–282.
[55] M. H. Goldstein, J. A. Schwade, and M. H. Bornstein, "The value of vocalizing: five-month-old infants associate their own noncry vocalizations with responses from caregivers," *Child Dev*, vol. 80, no. 3, pp. 636–644, 2009.
[56] A. N. Meltzoff, P. K. Kuhl, J. Movellan, and T. J. Sejnowski, "Foundations for a new science of learning," *Science*, vol. 325, no. 5938, pp. 284–288, 2009.

Omri Weissman received PhD in Psychology and Neuroscience from Bar-Ilan University in 2013. His dissertation project, entitled "Biological substrates of human affiliation" was mentored by Prof. Ruth Feldman. Specifically, Omri studied the involvement of the neuropeptide hormone Oxytocin (as well as other hormones pertaining human stress and affiliation) in social repertoire and parent-child interaction in a typically-developed sample. Currently, Omri conducts a post-doctoral-level research at Yale School of Medicine, where he studies the potential involvement of Oxytocin in chemoreception-related processes in human adults. Omri is motivated to pursue the quest concerning chemosensory phenomena and sensory-sensitivities, more so given that these processes are found to be disturbed (askew) in some psychiatric disorders, including Autism Spectrum Disorder. In the years to come, Omri anticipates to pursue a career as an independent investigator in the area of translational neuroscience.

Mohamed Chetouani is the head of the IMI2S (Interaction, Multimodal Integration and Social Signal) research group at the Institute for Intelligent Systems and Robotics (CNRS UMR 7222), University Pierre and Marie Curie-Paris 6. He received the M.S. degree in Robotics and Intelligent Systems from the UPMC, Paris, 2001. He received the PhD degree in Speech Signal Processing from the same university in 2004. In 2005, he was an invited Visiting Research Fellow at the Department of Computer Science and Mathematics of the University of Stirling (UK). Prof. Chetouani was also an invited researcher at the Signal Processing Group of Escola Universitaria Politecnica de Mataro, Barcelona (Spain). He is currently a Full Professor in Signal Processing, Pattern Recognition and Machine Learning at the UPMC. His research activities, performed at the Institute for Intelligent Systems and Robotics, cover the areas of social signal processing and personal robotics through non-linear signal processing, feature extraction, pattern classification and machine learning. He is also the co-chairman of the French Working Group on Human-Robots/Systems Interaction (GDR Robotique CNRS) and a Deputy Coordinator of the Topic Group on Natural Interaction with Social Robots (euRobotics). He is the Deputy Director of the Laboratory of Excellence SMART Human/Machine/Human Interactions In The Digital Society.

Catherine Saint-Georges received a MS degree in Psychopathology and Neurosciences from UPMC (Paris 6 University) and a MD from Paris XI University of Medicine in 2000. She specialized in child and adolescent psychiatry in 2001. She received the PhD degree in neuroscience from ED3C School (UPMC) in 2011. Her PhD topic dealt with dynamic of early interactions of autistic infants with their parents, with a special focus on synchrony, reciprocity and motherese. She is currently post-doc at the Institute of Intelligent Systems and Robotic (CNRS UMR 7222), where she is involved in research on autistic children and early mother-infant synchrony in infants at risk for autism. She is also Child Psychiatrist at la Salpetriere Hospital and in Andre Boulloche Center (Paris).

Nadège Bourvis received a MS degree in Biomathematics from Ecole Normale Supérieure de Paris and UPMC (Paris 6 University) in 2005 and a MD from Paris 6 University of Medicine in 2012. She is currently a PhD student in neuroscience from ED3C School (UPMC), working on social signal processing in stress situations in infants and adolescents, in both healthy and clinical populations. She is child psychiatrist in the Department of Child and Adolescent Psychiatry at Pitie-Salpetriere University Hospital in Paris.

Emilie Delaherche received the M.S. degree in Engineering from ESME Sudria (School of Engineering), Paris, 2005. She received the M.S.degree in Machine Learning and Signal Processing from the University Pierre and Marie Curie (UPMC), Paris, 2010. She received the PhD Degree in Computer Science from the same university in 2013. The topic of her thesis is the analysis of human-centered interaction dynamics. Since 2013, she is an Academic Technical Specialist at the MathWorks.

Orna Zagoory-Sharon received PhD in Biochemistry from Ben Gurion University in the Negev (Israel) in 2000. Over the past decade, Orna heads the ”wet lab” as an Associate Research Scientist at the Gonda Multidisciplinary Brain Sciences Center at Bar-Ilan University, where she works in close proximity with Prof. Ruth Feldman. Orna's interests surrounds hormonal systems (in humans) in relation to social and affiliative repertoire, and human development. She is especially interested in EIA/ELISA bio-essays for endocrine assessment, a topic on which she advise internationally about. Besides that, Orna holds a full position as a senior R&D officer in one of Israel's main health care providers.

David Cohen received a M.S. in neurosciences from the UPMC and the Ecole Normale Supérieure in 1987, and a M.D. from Necker School of Medicine in 1992. He specialized in child and adolescent psychiatry and certified in 1993. His first field of research was severe mood disorders in adolescent, topic of his PhD in neurosciences (2002). He is Professor at the UPMC and head of the department of Child and Adolescent Psychiatry at La Salpetriere hospital in Paris. His group runs research programs in the field of autism and other pervasive developmental disorder, severe mood disorder in adolescent and childhood onset schizophrenia and catatonia. He is a member of the International Association of Child and Adolescent Psychiatry and Allied Disciplines, the European College of Neuro-Psychopharmacology, the European Society of Child and Adolescent Psychiatry, and the Interna- tional Society of Adolescent Psychiatry.

Ruth Feldman is the Simms-Mann professor of psychology and neuroscience at Bar-Ilan University, Israel with a joint appointment at Yale University Medical School, Child Study Center. She is also the director of a community-based clinic for young children and their families and heads the Irving B. Harris internship program in early childhood clinical psychology. She received an MA in neuropsychology (1986) and clinical psychology (1988), PhD in developmental psychology (1994), and completed post-doctorate studies in developmental neuroscience (1996). Her research focuses on the biological basis of social affiliation, parent-infant relationship, bio-behavioral processes of emotion regulation, and she conducts several longitudinal studies from infancy to adulthood following infants born at high risk stemming from biological (prematurity), maternal (postpartum depression), and contextual (war-related trauma) risk conditions. She pioneered research on biobehavioral synchrony, the effects of touch interventions, and on the parental brain. Her research on the role of oxytocin in health and psychopathology has been instrumental to understanding the biological basis of social collaboration in humans.