UNCORRECTED FIRST PROOF

Acknowledgement: This is a draft of a chapter that has been accepted for publication by Oxford University Press in the forthcoming book *The Oxford Handbook of Emotional Development* edited by Dr. Daniel Dukes, Prof. Andrea C. Samson, and Prof. Eric A. Walle, and due for publication in 2021.

Early Interaction: New Approaches

Daniel S. Messinger[1]
Jacquelyn Moffitt[2]
Samantha G. Mitsven[2]
Amy Ahn[2]
Stephanie Custode[2]
Evgeniy Chervonenko[3]
Saad Sadiq[4]
Mei-Ling Shyu[4]
Lynn K. Perry[2]

[1]University of Miami
Department of Psychology
Department of Pediatrics
Department of Electrical and Computer Engineering
Department of Music Engineering
[2] University of Miami
Department of Psychology
[3] École Polytechnique fédérale de Lausanne
Department of Digital Humanities
[4] University of Miami
Department of Electrical and Computer Engineering

Corresponding Author:
Daniel S. Messinger (dmessinger@miami.edu)
University of Miami
Department of Psychology
5565 Ponce de Leon Blvd
Coral Gables, FL, 33146

Acknowledgements: Work on this chapter was supported by grants to the first and last author from the National Science Foundation (1620294, PI Messinger), the Institute of Education Sciences (R324A180203, PI Messinger), and the Miami Clinical and Translational Science Institute, from the National Center for Advancing Translational Sciences and the National Institute on Minority Health and Health Disparities (Co-PIs Messinger and Perry). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the funders.

Abstract

Early interaction is a dynamic, emotional process in which infants influence and are influenced by caregivers and peers. This chapter reviews new developments in behavior imaging—objective quantification of human action—and computational approaches to the study of early emotional interaction and development. Advances in the automated measurement and modeling of human emotional behavior—including objective measurement of facial expressions, machine learning approaches to detecting interaction and emotion, and electrophysiological measurements of emotional signals—provide new insights into how interaction occurs. Furthermore, advances in automated measurement and modeling can be applied to the study of atypical development, contributing to our understanding of, for example, social affective behaviors in toddlers with autism spectrum disorder (ASD). We conclude by posing questions for future directions of the field of computational approaches to emotion.

Keywords. Infants, machine learning, interaction, modeling, computational, electrophysiological, autism

Introduction

Early interaction between infants, parents and other caregivers is an emotional process replete with bouts of both laughter and distress. These emotional expressions often develop in the context of intricate social interactions that may be the basis of patterns of emotional engagement throughout the lifespan (Messinger, Ruvolo, Ekas, & Fogel, 2010). However, our understanding of emotional expression has been hampered because human coding of emotional expression is time-intensive (Cohn & Kanade, 2007). A consequence of this measurement bottleneck is that more is known about infants' perception of emotional expressions than of their actual production of these expressions (Mitsven, Messinger, Moffitt, & Ahn, in press). To surmount these difficulties, this chapter reviews computational approaches to the measurement and modeling of emotional expression and interaction. Modeling here refers both to advanced inferential (statistical) methods, machine learning approaches, and their increasingly common hybrids. Finally, we review recent work applying automated measurement of electrophysiological and behavioral indices of emotion to the characterization of autism spectrum disorder (ASD).

Automated measurement of emotional expression and interaction

Advances in machine learning (in which software learns to represent and classify video or audio signals) offer the possibility of automated measurement of facial expressions, emotional vocalizations, and other expressive actions. Here, we review three primary approaches to automated measurement of emotion. In the first approach, objective measures of low-level behavior features including the movement of facial landmarks and the proximity of infant and parent serve as direct indices of emotional functioning. In the second, unsupervised algorithms detect emotional signals directly from audio or video data. Here, the software detects and represents the phenomena of interest—and the human investigator interprets the results. The third and most common approach involves using algorithms to replicate human coding.

Low-level tracking methods

Tracking of emotional facial expressions. One approach to measuring emotional expressions, such as facial expressions, involves automated tracking of the movement of facial landmarks and head position in three-dimensional space from video (Jeni, Cohn, & Kanade, 2017). In an illustrative project, 13-month-olds were exposed to a positive (bubbles) and a negative (toy removal) emotion-eliciting task. Facial features exhibited greater displacement, velocity, and acceleration in response to the negative than the positive task, and infant head position showed the same pattern (Hammal et al., 2019). Together, the movement of facial features and head movement accounted for one third of the variance in manual behavioral affect ratings within each of the two conditions (Hammal, Cohn, Heike, & Speltz, 2015). Manual coding confirmed higher levels of smiles during positive tasks and higher levels of cry-faces (which encompass distress and anger expressions) during negative tasks (Hammal et al., 2018). The results suggest that low-level tracking of facial and head movement can distinguish negative (cry-face) versus positive (smiling) expressions.

Tracking movement and orientation. Low-level physical features of interaction have also been used to predict expert measurements of psychological constructs such as synchrony and mutual engagement. Leclère and colleagues (2016) combined 2D and 3D sensor data from 10 high-risk (referred for neglect) and 10 low-risk 1- to 3-year-olds and their mothers to examine mother-infant interactions during a pretend tea party. Kinect depth and video tracking indicated that higher levels of mother motion were associated with lower expert ratings of maternal sensitivity and intrusiveness, and higher ratings of infant avoidance. In addition, pauses in infant and parent joint movement were associated with higher ratings of maternal sensitivity and higher levels of infant engagement. The findings suggest that relatively low-level physical features such as mother-infant proximity and activity level are promising markers of caregiver sensitivity and intrusiveness and infant engagement, key indices of socioemotional development.

Unsupervised machine learning

A more radical approach to automated measurement involves direct unsupervised machine learning of emotional interaction from video or audio. Rehg and colleagues, for example, directly detected parent-child playful interaction characterized by quasi-periodic spatiotemporal patterns from posted YouTube videos (Prabhakar, Oh, Wang, Abowd, & Rehg, 2010). Likewise, Chu and colleagues (2017) automatically detected affective synchrony in videos of parents and infants engaged in face-to-face interaction. Using shape features of infant and mother faces, an unsupervised algorithm detected a priori areas of common action in overlapping segments of video that corresponded to infant and mother smile displays (see Figure 1). This is a bottom-up validation of the importance of positive emotion communication in early interaction. These approaches suggest the as yet unrealized potential of unsupervised machine learning to identify new patterns of early emotional interaction.

[Insert Messinger-Fig 1 here]

Computational approaches to replicate human coding

The most common approach to objective measurement is supervised training to replicate human expert measurements. One target is replication of the Facial Action Coding System (FACS; Ekman & Friesen, 1992; Ekman, Friesen, & Hager, 2002)—applied to infants in BabyFACS (Oster, 2006)—an expert system for documenting anatomically based appearance changes based on facial Action Units (Lucey, Ashraf, & Cohn, 2007; Mahoor et al., 2008). We previously instantiated automated measurement of the presence and intensity of Action Units by using nonlinear manifold learning (Belkin & Niyogi, 2003) of data by combining active appearance and shape models to train support vector machines (SVMs; Messinger, Mattson, Mahoor, & Cohn, 2012). This approach yielded insights into similarities between early positive and negative emotion expression, the structure of interactive positive affect, and early interaction dynamics.

Positive and negative expression similarities. Just as smiles are often used to index infant positive emotion, the cry-face is the preeminent infant expression of negative emotion. Importantly, both smiles and cry-face expressions can involve different degrees of mouth opening and Duchenne activation—eye constriction produced by the muscle orbiting the eyes. The Duchenne intensification hypothesis holds that Duchenne activation and mouth opening index the intensity of both smile and cry-face expressions (Bolzani-Dinehart et al., 2005; Darwin, 1872/1998). In support, both mouth opening and the Duchenne marker indexed greater perceived positive valence in smile expressions and greater perceived negative valence of cry-face expressions. Next, the intensification hypothesis was tested using the Face-to-Face/Still-Face (FFSF) protocol (Mattson, Cohn, Mahoor, Gangi, & Messinger, 2013, but see Mattson, Ekas, et al., 2013). In the FFSF, a naturalistic face-to-face interaction is interrupted when the parent is asked hold a still-face and not engage with the infant, and ends when the parent is asked to play again with the infant (Adamson & Frick, 2003; Tronick, Als, Adamson, Wise, & Brazelton, 1978). During face-to-face play, which is expected to elicit positive emotion, smiles were more likely to involve eye constriction than during the still-face, which elicits negative emotion (see Figure 2). As predicted, the proportion of cry-faces involving eye constriction during the negative emotion eliciting still-face was higher than during face-to-face play (Messinger et al., 2012). The results suggest that automated measurement of facial Action Units such as eye constriction can produce insights into the structure of infant positive and negative emotion expression.

[Insert Messinger-Fig 2 here]

Interactive positive affect. Use of the active appearance models described above (Mattson et al., 2013) to measure the Action Units involved in infant and parent smiling produced insights into the expression of positive emotion and the dynamic structure of early interaction. Some propose that only adult Duchenne smiling expresses positive emotion, whereas smiles without the Duchenne marker do not (Ekman & Friesen, 1982), although they do have other important social funcitons (see ~~Mireault, chapter XXX~~chapter MIREAULT, this volume). Objective measurement of the intensity of smiling and eye constriction in the face-to-face interactions of two dyads indicated that Duchenne smiling was not a discrete entity but a continuous signal (Messinger, Mahoor, Chow, & Cohn, 2009). Specifically, the intensity of smiling and eye constriction were highly correlated in both mothers and infants. In sum, neither infants nor mothers appeared to exhibit discrete Duchenne and non-Duchenne smiles during interaction (Messinger, Cassel, Acosta, Ambadar, & Cohn, 2008). Instead, all features of smiling covaried together, suggesting they indexed a continuum of positive emotion.

Interaction dynamics. Messinger et al. (2009) went on to describe early caregiver-infant interaction using a continuous measure of Duchenne smiling intensity derived from objective measurement of facial Action Unit intensity. This dynamic portrait of positive emotion uncovered variability in interactive synchrony at multiple temporal levels (see Figure 3). In Figure 3, changes in the zero-order correlation of infant and mother Duchenne smiling intensity illustrate variability in emotional synchrony over time. These changes suggest disruptions and repairs of emotional synchrony (Schore, 1994; Tronick & Cohn, 1989). Findings of dynamic changes in emotional synchrony are intriguing because a large body of research suggests that the degree to which parents adjust their own affective expressions to match those of their infants is associated with subsequent self-control, the internalization of social norms, and attachment security (Beebe et al., 2010; Kochanska, Forman, Aksan, & Dunbar, 2005; and see chapter HALBERSTADT, this volume).

[Insert Messinger-Fig 3 here]

Coding vocal expressions. In the audio domain, the use of physical characteristics to index emotional components of vocal expression is common. Bourvis and colleagues (2018) employed automated measures of infant and mother vocalization during the FFSF. These were supplemented with detection of an emotional component of mothers' speech, infant directed speech (e-IDS), indexed by higher pitch and wider pitch range. Infants increased their rate of vocalizing between the face-to-face and reunion episode of the FFSF but mothers exhibited few changes in vocalization parameters. In the reunion episode, likewise, infants increased their rate of response to mothers' e-IDS, rates of overlapping speech increased, and pauses in dyadic speech decreased. The results illustrate the potential of objective measures of the dyadic speech stream to disentangle patterns of emotional interaction following the still-face perturbation, a standard assessment of socioemotional functioning.

Coding attachment. Attachment security is central to early social and emotional development, and indexes an infant's ability to be comforted by a caregiver when distressed. Attachment security is typically assessed in the Strange Situation Procedure (SSP), which involves two brief separations from and reunions with the parent. However, attachment assessment is conventionally assessed using expert subjective ratings. Using relatively low-level, Kinect-based depth-video measurements of position and LENA-derived estimates of infant crying, Prince et al.

(2015) explored objectively measured attachment behavior in the reunion episodes of the SSP. Objective measurements of the frequency with which the infant made contact with the mother, the duration of that contact, the duration of infant crying and the inverse of the velocity of the infant's initial approach to the mother accounted for a substantial proportion of the variance in, respectively, expert ratings of proximity-seeking (approaching mother), contact-maintenance (staying close to mother), resistance (to contact with mother), and avoidance (ignoring or moving away from mother). These results suggest that measurement of physical proxemics and crying can provide insight into patterns of attachment previously captured exclusively via expert but subjective rating scales.

Chow and colleagues (2018) modeled "qualitative" changes in movement dynamics during the reunion episodes of the SSP by incorporating regime switching into a system of differential equations. Seeking a computational foundation for attachment theory, the researchers distinguished a proximity-seeking regime, in which infants tended to approach the parent, and an exploration regime, in which infants moved away from the parent to explore the room. As the infant attachment system became more activated in the second reunion, there was an increase in transitions to the proximity-seeking regime. These transitions were heightened in the presence of infant vocalizations (often cries), which functioned as signals of the infant's attachment needs. These results speak to an emerging capacity of researchers to computationally capture objectively measured infant- and dyad-specific emotional dynamics on a moment-to-moment basis to illuminate long-standing theories of early social motivation.

Modeling approaches to emotional expression and interaction

Computational approaches to the study of early emotion involve more than the use of machine learning algorithms to detect and measure expressive signals. Researchers are using increasingly sophisticated models to characterize when and why emotional signals are used during interaction, and to describe the development of those emotional interactions (for an advanced approach, see chapter ~~XXX: Rudrauf et al.~~RUDRAUF, this volume). Here we review research on the development of dyadic responses to infant distress, modeling of the predictability of smiling interactions, and the application of a novel framework for inferring infant goals during emotion-laden interactions.

Modeling face-to-face interactions

Chow et al. (2010) applied computational and statistical modeling approaches to understanding changes in infant and parent affective valence as they unfold in the FFSF. Specifically, a bivariate autoregressive model indicated the presence of both infant-to-parent and parent-to-infant interactive influence. Although each partner was responsive to the other, parents were more responsive to their infants than infants were to their parents. A stochastic regression approach applied within a multi-dyad time-series revealed changes in interactive influence over time that were accentuated in the reunion episode following the still-face. The results point to the importance of quantifying change over time to characterize how dyads respond to one another emotionally (Chow et al., 2014).

The goals of face-to-face interactions. Recently, our team used inverse optimal reinforcement modeling to infer likely infant and mother goals during their interactions (Ruvolo, Messinger, & Movellan, 2015). Probable consequences of beginning and ending smiles on the durations of subsequent dyadic states such as mutual smiling were used to infer goals. Results of this modeling approach suggest that mothers' likely goal is to increase the duration of mutual smiling (see Figure 4). However, infants' likely goal is to increase the duration of epochs when mother is smiling but the infant is not. To achieve this goal, infants briefly smile until the mother smiles, and then they end their own smile. These results are surprising as they suggest infants do not act to increase the time they express positive emotion. Instead, infants smile as part of a dyadic process in which they create and then disengage from moments of mutual positive emotion expression (Stifter & Moyer, 1991).

[Insert Messinger-Fig 4 here]

Development changes in face-to-face interactions. We examined the predictability of infants initiating or ending a smile within particular face-to-face interactive contexts observed weekly from 1 to 6 months of age (Messinger, Ruvolo, Ekas, & Fogel, 2010). The mean, variance, and overall distribution of mutual smiling states became more similar over consecutive weekly sessions with age, such that individual dyads' states of mutual positive affect became more predictable—to each partner, as well as to an outside observer—with development. Infants and mothers also increased the number of alternating turns in turn-taking interactions involving initiating and terminating smiles, suggesting that infants and mothers became more emotionally responsive to one another with age (Messinger et al., 2010). These findings suggest that repeated infant-parent interactions produce stable dyadic differences in emotional expressivity.

Developmental consequences of face-to-face interactions. Ekas, Haltigan, and Messinger (2013) examined continuous trends in manually coded infant expressivity over the course of the still-face using multilevel models (see Figure 5). Group effects indicated logarithmic decreases in infant gazing at the parent and smiling and increases in infant cry-face expressions. At the level of individual trajectories, infant gaze (but not smiling) trajectories were associated with later attachment security in a theoretically meaningful fashion (Ainsworth, Blehar, Waters, & Wall, 1978). Infants with later insecure-avoidant attachment exhibited the steepest drop in gazing at the parent (disengagement with the attachment figure), infants with later insecure-resistant attachment exhibited the least drop in gazing (they remained engaged with the parent despite their unavailability), and securely attached infants exhibited a moderate slope of disengagement. The results suggest that dynamic modeling of changes in engagement over time during the negative emotion eliciting still-face may be associated with later patterns of socioemotional security.

[Insert Messinger-Fig 5 here]

Modeling naturally occurring elicitors of emotional interactions

Researchers have combined computational modeling (e.g., Hidden Markov Models, or HMMs) and statistical (e.g., cluster analysis) approaches to understanding infant-mother interaction in natural contexts—in this case, dyadic responses to childhood inoculations (Backer, Quigley, & Stifter, 2018; Stifter & Rovine, 2015). Studies investigating interactive processes involved in the down-regulation of infant distress following immunization have traditionally relied on correlational or contingency analyses to understand the effectiveness of maternal soothing behaviors on infant distress. However, such approaches are unable to capture the influence of multiple simultaneous soothing behaviors that occur in response to infant distress. HMMs indicated that infants utilized more complex responses to aversive stimuli and became more organized and efficient in their soothing behaviors with age (Stifter & Rovine, 2015). Cluster analyses indicated that the fit between infants' capacity to be soothed (indexed by temperamental factors) and appropriate and responsive changes in maternal soothing behaviors over time that determined infant soothability. These findings suggest the potential of an integrative approach to modeling the reciprocal interplay of emotional communication between parent and child over time (Backer et al., 2018).

Modeling emotional vocalizations

Infant cries are a central focus of automated measurement research on emotional components of the vocal signal. Infant crying is a universal distress signal that becomes a more heterogenous negative emotion expression over the first year (Gustafson & Green, 1991). The commercially available Language ENvironment Analysis (LENA) technology employs Gaussian mixture models to detect adult speech, infant speech and emotion-laden non-speech vocalizations, which tend to be cries and are referred to as such here.

The temporal and interactive dynamics of crying. In day-long home recordings, Fields-Olivieri and Cole (2019) found that mothers were less likely to respond to toddlers' cries than toddler's word-like vocalizations. But when mothers did respond to toddlers' cries, the toddlers were more likely to subsequently produce speech-like vocalizations rather than additional cries (Fields-Olivieri & Cole, 2019). With respect to temporal structure, Abney and colleagues (2017) found that home-recorded cries in the first year exhibited a higher degree of clustering in time (temporal heterogeneity) than speech-like vocalizations. Likewise, among 1- to 2-year-olds in an early intervention preschool classroom, we found that vocal expressions of negative affect perpetuated themselves in time (the duration of one cry predicted the duration of the next) and cries tended to occur in clusters over the day (burstiness; Messinger et al., 2019). Together these results highlight the power of objective measurement of cries to shed light on the temporal structure of negative affect and the dynamics of early communication using day-long samples of naturally occurring behavior.

Using automated measurement and modeling to understand atypical development

Researchers have begun using automated measurement, including electrophysiological approaches, to measure individual differences in children with autism spectrum disorder (ASD). ASD is a pervasive disorder of social communication that impacts both nonverbal and verbal interaction (APA, 2013; see chapter CONNER, this volume). We begin by describing electrophysiological measurement of arousal and then review its application to ASD. We then review work using machine learning of behavior to index ASD symptoms during diagnostic assessments.

Tracking Arousal. Physiological indices of arousal are a key index of emotional dynamics. Electrodermal activity (EDA) measured by skin conductance, for example, can index sympathetic nervous system (SNS) arousal, providing a physiologic indicator of children's emotional responses and regulation (Benedek & Kaernbach, 2010; Chow et al., 2010; Rogers & Ozonoff, 2005). Measurement of EDA captures the SNS "fight or flight" response and considers both the slow-changing levels of arousal (tonic EDA) and immediate responses to the environment (phasic EDA; Fowles, 2007). Phasic changes in EDA are the result of fluctuations in eccrine sweat function in response to sympathetic activation (Fowles, 2007). EDA is widely used as an indicator of emotional arousal (Bouscein, 2012). In neonates, noxious stimuli—including a heel prick procedure (Harrison et al., 2006) and high sound levels (Salavitabar et al., 2010)—have been tied to sharp, sustained increases in EDA. By contrast, cessation of nursing is associated with a reduction in EDA below baseline levels (Harrison et al., 2006).

Electrodermal activity in children with ASD. Recent technological developments have enabled ambulatory measurement of EDA via wearable wrist sensors approximately the size and appearance of a watch (Poh et al., 2012; Poh, Swenson, & Picard, 2010). These ambulatory measurements provide a unique understanding of individual differences in response to environmental stimuli and interactions. In a sample of children with ASD (4-10 years), the concordance of ambulatory measures of parent and child EDA during a free-play period was lower in dyads in which the child had higher autism symptoms (Baker et al., 2015). Over developmental time, it is possible that autism-related social impairment interrupts the development of synchronous interactions between child and parent. Toddlers with ASD with higher restricted and repetitive behavior scores on the Autism Diagnostic Observation Schedule or ADOS-2 (Lord et al., 2012), the gold standard, play-based assessment of ASD, have greater increases in skin conductance level (SCL) in response to mechanical toys as opposed to passive toys (Prince et al., 2017). This lends credence to the idea that children with higher autism symptoms are differentially reactive to specific stimuli in the immediate environment in a way that may preclude concordance with the parent. In both children with typical development and children with ASD, low EDA appears to be a risk factor for externalizing behavior problems in the context of harsh or low-quality parenting (Baker et al., 2017; El-Sheikh & Erath, 2011). Strikingly, instances of severe physical aggression for inpatient, minimally verbal, school age children with ASD can be predicted one minute ahead based on ambulatory monitoring of sympathetic (EDA) and parasympathetic (cardiac) arousal (Goodwin et al., 2018). The ambulatory measurement of arousal is a promising tool for understanding individual differences in how children with and without ASD interface with their social and physical environments.

Measuring ASD symptoms with machine learning. During the ADOS-2 assessment of ASD, a trained clinical examiner assesses autism symptoms. We were interested in predicting ADOS-2 social affect symptoms, which index deficits in the quantity and quality of vocal initiations, gesturing, and facial expressions including smiles, as well as unusual eye contact. Processing video with the Affdex system (Stockli, Schulte-Mecklenbeck, Borer, & Samson, 2018), objective measurements of social smiling to the examiner and parent from video were inversely associated with ADOS social affect symptoms (Ahn et al., 2019; Moffitt et al., 2019). LENA measures of adult-child turn-taking during the ADOS were also moderately associated with social affect symptoms such that higher turn-taking was associated with lower symptom levels. We next used deep learning to directly predict social affect symptoms from the ADOS-2 audio stream (Sadiq et al., 2019). Deep learning algorithms take raw data as input and represent features of these data in sequential layers whose output can be a classification (Bishop, 2006; LeCun, Bengio, & Hinton, 2015) of audio or video signals (Lavner, Cohen, Ruinskiy, & Ijzerman, 2016). We combined neural networks with recurrence and memory features to leverage temporal sequencing with a Synthetic Random Forest, a nonlinear algorithm in which the sequential interplay of input features that correspond to the branches of virtual trees, predict outcomes (Lu, Sadiq, Feaster, & Ishwaran, 2018). This deep learning approach predicted social affect severity scores more effectively than the pre-trained LENA algorithm (Sadiq et al., 2019). Together, the results highlight the potential of different forms of machine learning to directly estimate emotional symptoms in children being assessed for ASD (Hashemi et al., 2017).

Conclusions

Infants' early interaction and emotional expressions set the stage for emotional functioning throughout the lifespan. Objective measurement of behavior and computational modeling are providing insights into how infants express emotion, and how emotional interactions unfold in real time and over development. Applications of these approaches to children with ASD suggest the potential utility of objective measurement of the emotional component of autism symptoms, and the role of psychophysiological measurements of arousal in understanding individual differences in children with ASD.

Future Directions

Objective measurement of children's emotional behavior by means of deep learning is in its infancy. The synthesis of multimodal emotional parameters (e.g., facial, vocal, and movement) remains an important goal, as does the integration of these objective measurements with psychophysiological indices of constructs such as arousal. Likewise, the ability of automated measurement to facilitate studies of children’s emotional functioning over substantial periods of time and multiple contexts (e.g., home, preschool, and clinic) remains a goal, as does the objective study of children’s emotional interactions with peers as well as parents. Finally, computational modeling of emotional interaction is increasing in its ability to understand moment-to-moment changes in affective states. However, modeling of objective measurement to better understand emotional development remains aspirational.

Figures

Figure 1. Discovered synchronies in six parent-infant dyads. Strong smiles and mutual attention were among the synchronies discovered between parents and their 6-month-old infants. Credit: Chu, W.-S., De la Torre, F., Cohn, J. F., & Messinger, D. S. (2017). A branch-and-bound framework for unsupervised common event discovery. *International journal of computer vision*, *123*(3), 372-391. doi:10.1007/s11263-017-0989-7

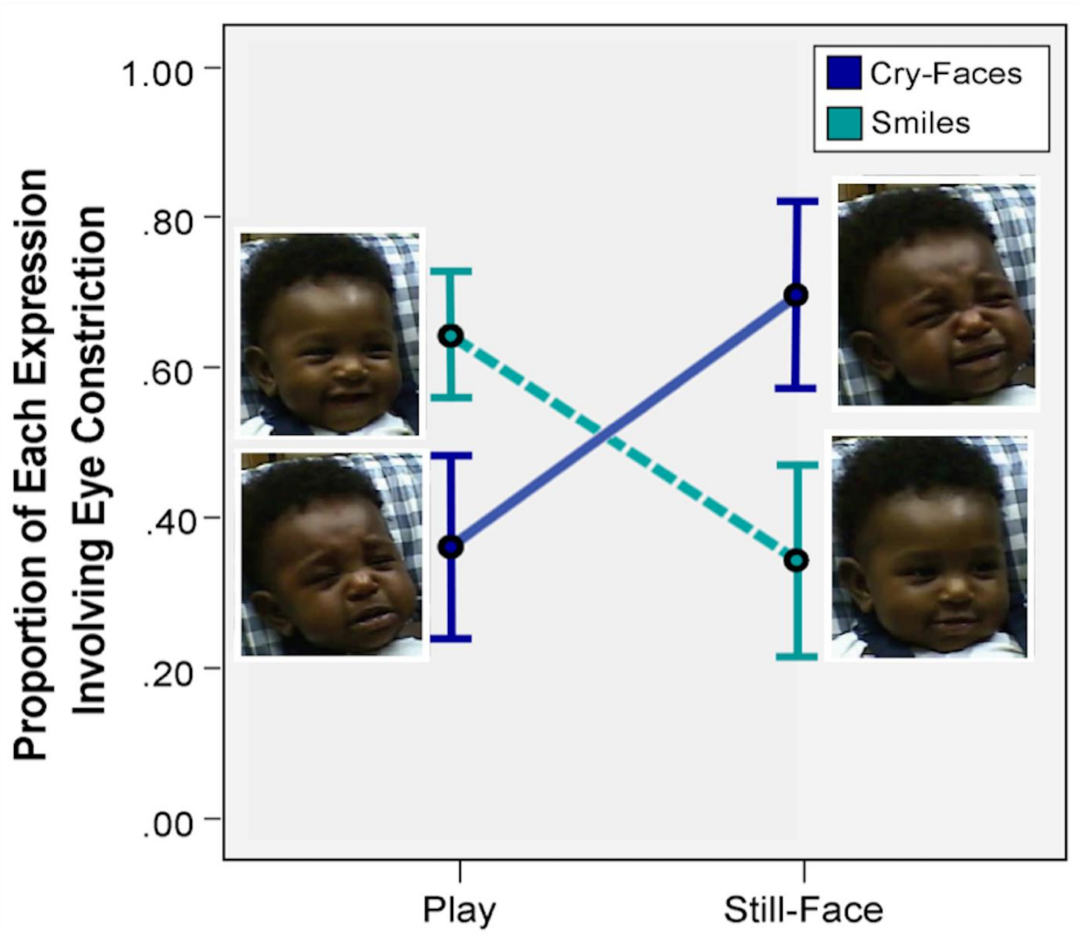

Figure 2. Eye constriction (the Duchenne marker) indexes positive and negative affective intensity in the Face-to-Face/Still-Face (FFSF). Smiling during the face-to-face play with the parent involved a higher proportion of smiling with eye constriction than smiling during the still-face. The still-face involved a higher proportion of cry-faces with eye constriction than face-to-face play. Credit: Mattson, W. I., Cohn, J. F., Mahoor, M. H., Gangi, D. N., & Messinger, D. S. (2013). Darwin's Duchenne: Eye constriction during infant joy and distress. *PloS One, 8*(11), e80161.

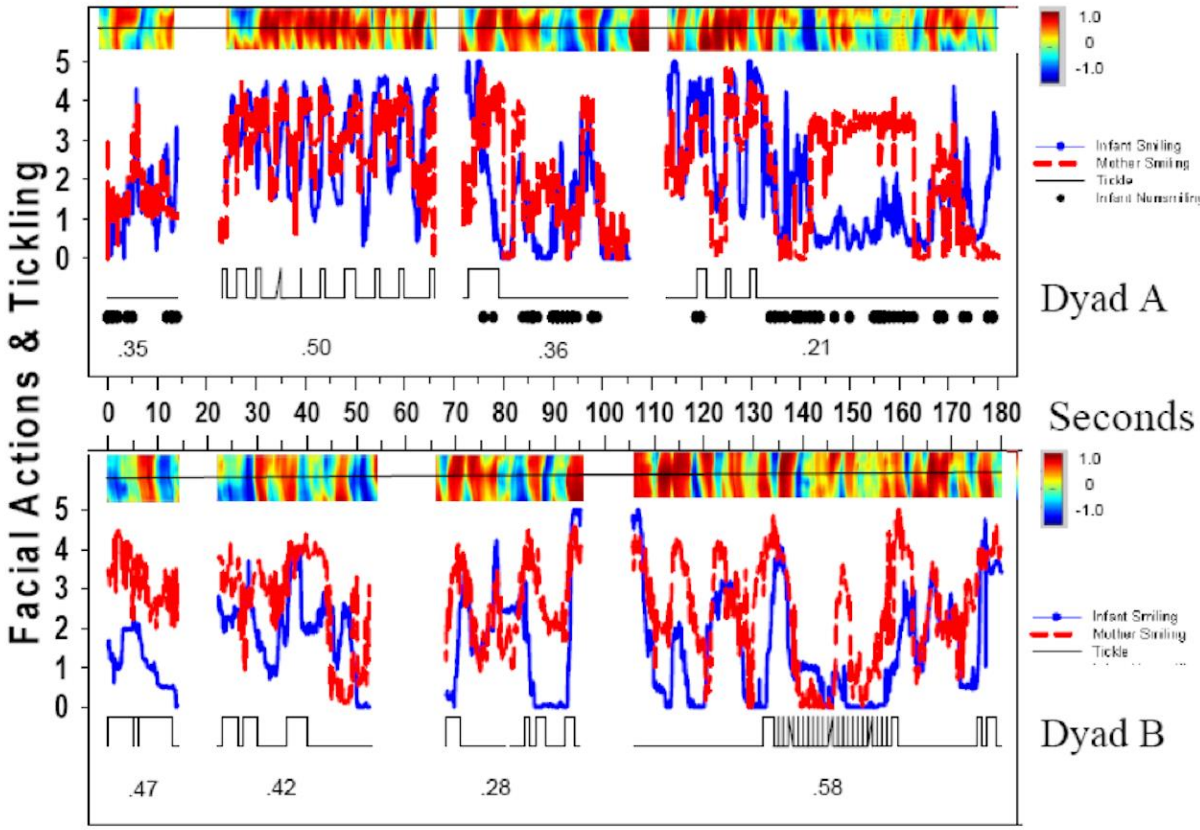

Figure 3. Automated Interaction. Correlations between infant and mother smiling activity are displayed. Above each segment of interaction is a plot of the windowed cross-correlations for successive three second segments of interaction. High positive correlations are deep red and high negative correlations are deep blue (see color bar at right). The horizontal midline of the plots indicates the zero-order correlation; lagged correlations are indicated above and below the midline. Credit: Messinger, D. S., Mahoor, M. H., Chow, S.-M., & Cohn, J. F. (2009). Automated measurement of facial expression in infant–mother interaction: A pilot study. *Infancy, 14*(3), 285-305. doi:10.1080/15250000902839963

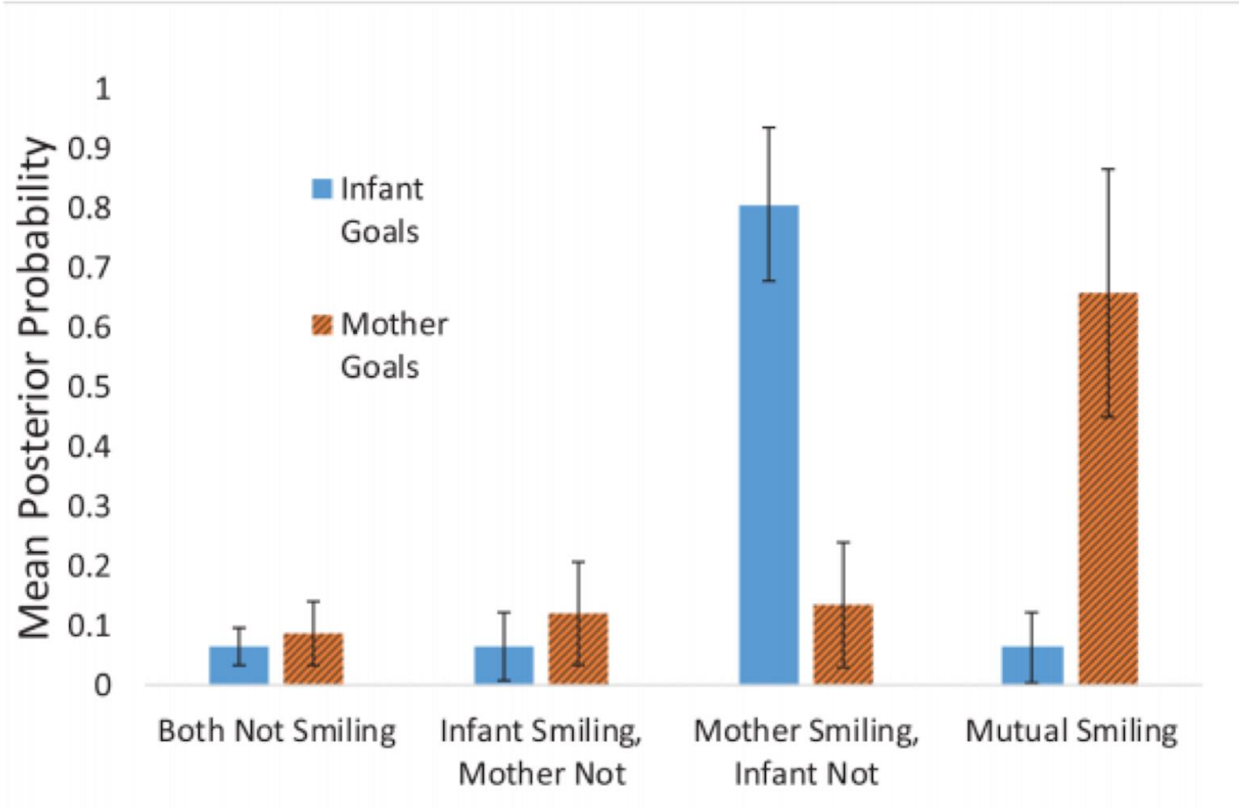

Figure 4. Means of the probability distributions of potential mother and infant goals. Error bars are 95% confidence intervals of the mean. Credit: Ruvolo, P., Messinger, D., & Movellan, J. (2015). Infants time their smiles to make their moms smile. PloS One, *10(*9), e0136492. doi:10.1371/journal.pone.0136492

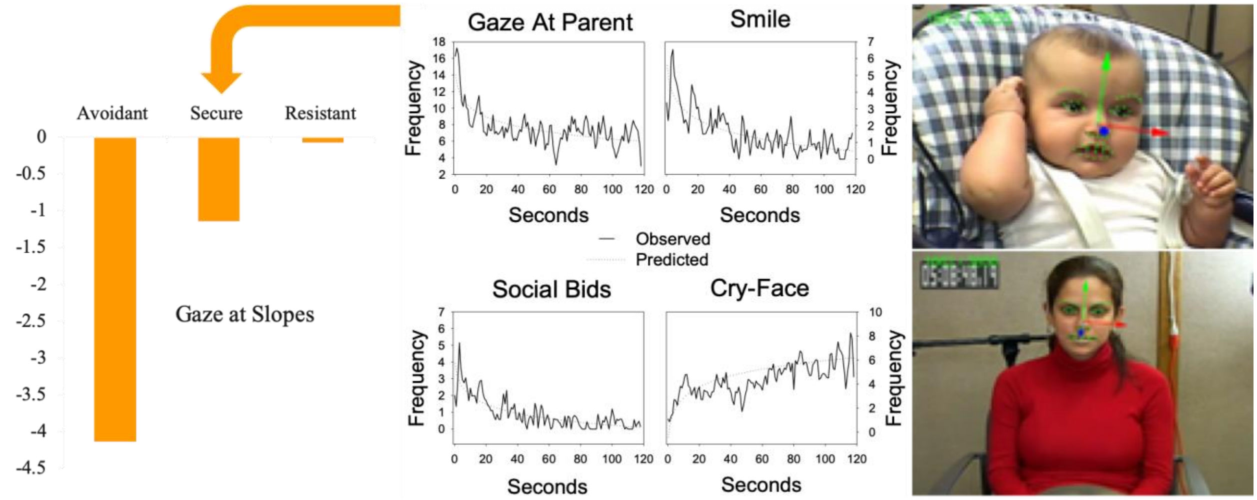

Figure 5. Observed and predicted mean frequencies of A) Gazes at parent, B) Smiles, C) Positive social bids, and D) Cry-face expressions over time in the still-face episode. Frequencies refer to the number of frames per second (maximum 30) in which a particular behavior occurred. Social bids were defined as smiles in the presence of gazing at the parent. Predicted refers to the expected frequency based on a hierarchical linear model containing an intercept and a linear term indexing behavior change proportional to log10 transformation of the number of seconds elapsed. Although the model only contains linear terms, the log transformation allows for curvilinear change over seconds. Credit: Ekas, N. V., Haltigan, J. D., & Messinger, D. S. (2013). The dynamic still-face effect: Do infants decrease bidding over time when parents are not responsive? *Developmental Psychology*, *49*(6), 1027-1035. doi:10.1037/a0029330.

References

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2017). Multiple Coordination Patterns in Infant and Adult Vocalizations. *Infancy, 22*(4), 514-539. doi:doi:10.1111/infa.12165

Adamson, L. B., & Frick, J. E. (2003). The still face: A history of a shared experimental paradigm. *Infancy, 4*, 451-473.

Ahn, Y.A., Moffitt, J., Durocher, J., Hale, M., & Messinger, D. (2019, May) *Objectively measured social communication behaviors during the ADOS-2.* Poster presented at 2019 Meeting of the International Society for Autism Research (INSAR). Montreal, QC.

Ainsworth, M. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Lawrence Erlbaum.

APA. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM 5).* Washington, DC: American Psychiatric Association.

Backer, P. M., Quigley, K. M., & Stifter, C. A. (2018). Typologies of dyadic mother-infant emotion regulation following immunization. *Infant Behavior and Development*, *53*, 5-17. doi:https://doi.org/10.1016/j.infbeh.2018.09.007

Baker, J. K., Fenning, R. M., Erath, S. A., Baucom, B. R., Moffitt, J., & Howland, M. A. (2017). Sympathetic Under-Arousal and Externalizing Behavior Problems in Children with Autism Spectrum Disorder. *J Abnorm Child Psychol*. doi:10.1007/s10802-017-0332-3

Baker, J. K., Fenning, R. M., Howland, M. A., Baucom, B. R., Moffitt, J., & Erath, S. A. (2015). Brief Report: A Pilot Study of Parent-Child Biobehavioral Synchrony in Autism Spectrum Disorder. *J Autism Dev Disord, 45*(12), 4140-4146. doi:10.1007/s10803-015-2528-0

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., . . . Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & Human Development, 12*(1), 3 - 141.

Belkin, M., & Niyogi, P. (2003). Laplacian Eigenmaps for dimensionality reduction and data representation. *Neural Computation Archive, 15*(6), 1373 - 1396.

Benedek, M., & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology, 47*(4), 647-658.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer.

Bolzani-Dinehart, L., Messinger, D. S., Acosta, S., Cassel, T., Ambadar, Z., & Cohn, J. (2005). Adult perceptions of positive and negative infant emotional expressions. *Infancy, 8*(3), 279–303.

Bourvis, N., Singer, M., Saint Georges, C., Bodeau, N., Chetouani, M., Cohen, D., & Feldman, R. (2018). Pre-linguistic infants employ complex communicative loops to engage mothers in social exchanges and repair interaction ruptures. *Royal Society Open Science, 5*: 170274. https://doi.org/10.1098/rsos.170274

Bouscein, W. (2012). *Electrodermal Activity* (2 ed.). New York: Springer.

Calvo, R. A. (2010). Latent and emergent models in affective computing. *Emotion Review, 2*(3), 288-289. doi:10.1177/1754073910368735

Chow, S. M., Ou, L., Ciptadi, A., Prince, E. B., You, D., Hunter, M. D., . . . Messinger, D. S. (2018). Representing Sudden Shifts in Intensive Dyadic Interaction Data Using Differential Equation Models with Regime Switching. *Psychometrika*. doi:10.1007/s11336-018-9605-1

Chow, S., Haltigan, J. D., & Messinger, D. S. (2010). Dynamic infant-parent affect coupling during the Face-to-Face/Still-Face. *Emotion*, *10*, 101-114.

Chow, S.-M., Mattson, W. I., & Messinger, D. S. (2014). Representing Trends and Moment-to-Moment Variability in Dyadic and Family Processes Using State-Space Modeling Techniques. In S. M. McHale, P. Amato, & A. Booth (Eds.), *Emerging Methods in Family Research* (Vol. 4, pp. 39-55): Springer International Publishing.

Chu, W.-S., De la Torre, F., Cohn, J. F., & Messinger, D. S. (2017). A Branch-and-Bound Framework for Unsupervised Common Event Discovery. *International Journal of Computer Vision*, 1-20. doi:10.1007/s11263-017-0989-7

Cohn, J., & Kanade, T. (2007). Automated facial image analysis for measurement of emotion expression. In J. A. Coan & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment* (pp. 222-238). New York: Oxford.

Darwin, C. (1872/1998). *The expression of the emotions in man and animals (3rd edition)*. New York: Oxford University.

Ekas, N. V., Haltigan, J. D., & Messinger, D. S. (2013). The dynamic still-face effect: do infants decrease bidding over time when parents are not responsive? *Developmental Psychology*, *49*(6), 1027-1035. doi:10.1037/a0029330

Ekman, P., & Friesen, W. (1992). *Changes in FACS Scoring (Instruction Manual)*.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior*, *6*(4), 238-252.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System Investigator's Guide*.

El-Sheikh, M., & Erath, S. A. (2011). Family conflict, autonomic nervous system functioning, and child adaptation: State of the science and future directions. *Development and Psychopathology, 23*(2), 703-721. doi:10.1017/S0954579411000034

Fields-Olivieri, M. A., & Cole, P. M. (2019). Sequences of toddler negative emotion and parent–toddler verbal communication during a waking day. *Infancy, 24*(6), 857-880. doi:10.1111/infa.12310

Fowles, D. C., & Fowles, D. C. (2007). *The Measurement of Electrodermal Activity in Children*

Goodwin, M.S., Oxdenizci, O., Cumpanasoiu, C., Tian, P., Guo, Y., Stedman, A., Peura, C., Mazefsky, C., Siegel, M., Erdogmus, D., & Ioannidis, S. (2018). Predicting imminent aggression onsetn in minimally-verbal youth with autism spectrum disorder using preceding physiological signals. *International Conference on Pervasive Computing Technologies for Healthcare,* 2018 May, 201-207. 10.1145/3240925.3240980

Gustafson, G. E., & Green, J. A. (1991). Developmental coordination of cry sounds with visual regard and gestures. *Infant Behavior & Development, 14*(1), 51-57. doi:10.1016/0163-6383(91)90054-V

Hammal, Z., Cohn, J. F., Heike, C., & Speltz, M. L. (2015). Automatic Measurement of Head and Facial Movement for Analysis and Detection of Infants' Positive and Negative Affect. *Frontiers in ICT, 2*(21). doi:10.3389/fict.2015.00021

Hammal, Z., Cohn, J. F., Wallace, E. R., Heike, C. L., Birgfeld, C. B., Oster, H., & Speltz, M. L. (2018). Facial Expressiveness in Infants With and Without Craniofacial Microsomia: Preliminary Findings. *The Cleft Palate-Craniofacial Journal*, 1055665617753481.

Hammal, Z., Wallace, E. R., Heike, C. L., & Speltz, M. L. (2017). *Automatic AU detection in infants using convolutional neural network*. Paper presented at the Proceedings of the

International Conference on Affective Computing and Intelligent Interaction,, San Antonio, TX.

Hammal, Z., Wallace, E. R., Speltz, M. L., Heike, C. L., Birgfeld, C. B., & Cohn, J. F. (2019). Dynamics of Face and Head Movement in Infants with and without Craniofacial Microsomia: An Automatic Approach. *Plastic and Reconstructive Surgery – Global Open*, *7*(1), e2081. doi:10.1097/gox.0000000000002081

Harrison, D., Boyce, S., Loughnan, P., Dargaville, P., Storm, H., & Johnson, L. (2006). Skin conductance as a measure of pain and stress in hospitalised infants. *Early Human Development*, *82*, 603-608. https://doi.org/10.1016/j.earlhumdev.2005.12.008

Hashemi, J., Spina, T.V., Tepper, M., Esler, A., Morellas, V., Papanikolopoulos, N., & Sapiro, G. (2012). A computer vision approach for the assessment of autism-related behavioral markers. Paper presented at the 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics. doi: 10.1109/DevLrn.2012.6400865

Jeni, L., Cohn, J.F., & Kanade, T. (2017). Dense 3D face alignment from 2D video for real-time use. *Image and Vision Computing*, *58*, 13-24. https://doi.org/10.1016/j.imavis.2016.05.009

Kochanska, G., Forman, D. R., Aksan, N., & Dunbar, S. B. (2005). Pathways to conscience: early mother-child mutually responsive orientation and children's moral emotion, conduct, and cognition. *Journal of Child Psychology and Psychiatry*, *46*(1), 19-34. doi:10.1111/j.1469-7610.2004.00348.x

Lavner, Y., Cohen, R., Ruinskiy, D., & Ijzerman, H. (2016). Baby Cry Detection in Domestic Environment using Deep Learning. *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*. doi:10.1109/ICSEE.2016.7806117

Leclère, C., Avril, M., Viaux-Savelon, S., Bodeau, N., Achard, C., Missonnier, S., . . . Cohen, D. (2016). Interaction and behaviour imaging: a novel method to measure mother–infant interaction using video 3D reconstruction. *Translational Psychiatry, 6*, e816. doi:10.1038/tp.2016.82 https://www.nature.com/articles/tp201682#supplementary-information

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436. doi:10.1038/nature14539

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., Bishop, S. L., & (Firm), W. P. S. (2012). *ADOS-2: Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.

Lu, M., Sadiq, S., Feaster, D., & Iswaran, H. (2018). Estimating Individual Treatment Effect in Observational Data Using Random Forest Methods. *Journal of Computational and Graphical Statistics, 27*(1), 209-219. doi: 10.1080/10618600.2017.1356325

Lucey, S., Ashraf, A. B., & Cohn, J. (2007). Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face. In K. Kurihara (Ed.), *Face Recognition*. Mammendorf, Germany: Pro Literatur Verlag.

Mahoor, M. H., Messinger, D. S., Ibanez, L., Kimijima M., Wang, Y., Cadavid, S., & Cohn, J. F. (2008). *Studying Facial Expressions Using Manifold Learning and Support Vector Machines.* Paper presented at the IEEE 7th International Conference on Development and Learning, Monterey, CA.

Mattson, W. I., Cohn, J. F., Mahoor, M. H., Gangi, D. N., & Messinger, D. S. (2013). Darwin’s Duchenne: Eye Constriction during Infant Joy and Distress. *PLoS ONE, 8*(11), e80161.

Mazefsky, C. A., Pelphrey, K. A., & Dahl, R. E. (2012). The need for a broader approach to emotion regulation research in autism. *Child Development Perspectives*, *6*(1), 92-97. doi:10.1111/j.1750-8606.2011.00229.x

Messinger, D. S., Custode, S., Perry, L. K., Mitsven, S. G., Ullery, M., Katz, L. F., Viatle, L., Valtierra, A. M., Prince, E. B., & Johnson, N. (2019). *Classroom crying: Temporal characteristics at one year.* Oral Presentation at the biennial meeting of the Society for Research in Child Development. Baltimore, MD.

Messinger, D. S., Mahoor, M. H., Chow, S.-M., & Cohn, J. F. (2009). Automated measurement of facial expression in infant–mother interaction: A pilot study. *Infancy*, *14*(3), 285-305. doi:10.1080/15250000902839963

Messinger, D. S., Mattson, W. I., Mahoor, M. H., & Cohn, J. F. (2012). The eyes have it: making positive expressions more positive and negative expressions more negative. *Emotion, 12*(3), 430-436.

Messinger, D., Cassel, T., Acosta, S., Ambadar, Z., & Cohn, J. (2008). Infant Smiling Dynamics and Perceived Positive Emotion. *Journal of Nonverbal Behavior*, *32*, 133-155.

Messinger, D., Mahoor, M., Cadavid, S., Kimijima, M., Haltigan, J. D., & Cohn, J. (2009). *The Role of Eye Constriction in Positive and Negative Infant Emotional Expressions.* Paper presented at the International Society for Research in Emotions, Leuven, Belgium.

Messinger, D., Ruvolo, P., Ekas, N., & Fogel, A. (2010). Applying Machine Learning to Infant Interaction: The Development is in the Details. *Neural Networks*, *23*(10), 1004–1016.

Mitsven, S., Messinger, D.S., Moffitt, J., & Ahn, Y.A. (In Press). Infant emotional development. In C. Tamis-Lemonda & J.J. Lockman (Eds.) *Handbook on Infant Development.* Cambridge, United Kingdom: Cambridge University Press.

Moffitt, J., Tao, Y., Ahn, Y.A., Sadiq, S., Shyu, M., & Messinger, D.S. (2019, May) *Objective measurement of movement during ADOS-2 assessment for young children with ASD.* Poster presented at the 2019 Meeting of the Social and Affective Neuroscience Society (SANS). Miami, FL.

Montirosso, R., Riccardi, B., Molteni, E., Borgatti, R., & Reni, G. (2010). Infant's emotional variability associated to interactive stressful situation: A novel analysis approach with Sample Entropy and Lempel–Ziv Complexity. *Infant Behavior and Development, 33*(3), 346-356. doi:http://dx.doi.org/10.1016/j.infbeh.2010.04.007

Oster, H. (2006). Baby FACS: Facial Action Coding System for Infants and Young Children. *Unpublished monograph and coding manual. New York University*.

Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Trans Biomed Eng, 57*(5), 1243-1252. doi:10.1109/tbme.2009.2038487

Poh, M.-Z., Loddenkemper, T., Reinsberger, C., Swenson, N. C., Goyal, S., Sabtala, M. C., . . . Picard, R. W. (2012). Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor. *Epilepsia, 53*(5), e93-e97. doi:10.1111/j.1528-1167.2012.03444.x

Prabhakar, K., Oh, S., Wang, P., Abowd, G. D., & Rehg, J. M. (2010). *Temporal causality for the analysis of visual events.* Paper presented at the Computer Vision and Pattern Recognition (CVPR), IEEE Conference.

Prince, E. B., Ciptadi, A., Gangi, D., Martin, K., Rozga, A., Rongfang, J., Rehg, J., and Messinger, D. (2015). *Automated Measurement of Dyadic Interaction Predicts Expert*

Ratings of Attachment in the Strange Situation. Paper presented at the Association for Psychological Science Annual Convention, New York, New York.

Prince, E. B., Kim, E. S., Wall, C. A., Gisin, E., Goodwin, M. S., Simmons, E. S., . . . Shic, F. (2017). The relationship between autism symptoms and arousal level in toddlers with autism spectrum disorder, as measured by electrodermal activity. *Autism, 21*(4), 504-508. doi:10.1177/1362361316648816

Rehg, J. M., Abowd, G. D., Rozga, A., Romero, M., Clements, M. A., Sclaroff, S., . . . Kim, C. (2013). *Decoding children's social behavior.* Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Rogers, S.J., & Ozonoff, S. (2005). Annotation: What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 46*(12), 1255-1268. doi: 10.1111/j.1469 10.1111/j.1469-7610.2005.01431.x-7610.2005.01431.x

Ruvolo, P., Messinger, D., & Movellan, J. (2015). Infants Time Their Smiles to Make Their Moms Smile. *PLoS ONE, 10*(9), e0136492. doi:10.1371/journal.pone.0136492

Sadiq, S., Castellanos, M. A., Moffitt, J. M., Shyu, M. L., Perry, L. K., & Messinger, D. S. (2019). *Deep Learning based Multimedia Data Mining for Autism Spectrum Disorder (ASD) Diagnosis*. Paper presented at the Translational Multimedia Data Mining (TMDM) workshop at the IEEE International Conference on Data Mining, Beijing, China.

Salavitabar, A., Haidet, K.K., Adkins, C.S., Susman, E.J., Palmer, C., Storm, H. (2010). Preterm infants' sympathetic arousal and associated behavioral responses to sound stimuli in the Neotatal Intensive Care Unit. *Advances in Neonatal Care, 10,* 158-166. 10.1097/ANC.0b013e3181dd6dea

Schore, A. N. (1994). *Affect Regulation & the Origin of Self: The Neurobiology of Emotional Development*. Hillsdale, NJ: Erlbaum.

Stockli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial expression analysis with AFFDEX and FACET: A validation study. *Behav Res Methods, 50*(4), 1446-1460. doi:10.3758/s13428-017-0996-1

Stifter, C. A., & Moyer, D. (1991). The regulation of positive affect: Gaze aversion activity during mother-infant interaction. *Infant Behavior and Development, 14*, 111-123.

Stifter, C. A., & Rovine, M. (2015). Modeling dyadic processes using hidden Markov models: A time series approach to mother–infant interactions during infant immunization. . *Infant and Child Development, 24*(3), 298-321. doi:10.1002/icd.1907

Torres, R., Battaglino, D., & Lepauloux, L. (2017). Baby Cry Sound Detection: A Comparison of Hand Crafted Features and Deep Learning Approach. In B. G., I. L., J. C., & L. A. (Eds.), *Engineering Applications of Neural Networks. EANN 2017. Communications in Computer and Information Science, vol 744.* New York, New York: Springer Nature.

Tronick, E. Z., & Cohn, J. F. (1989). Infant-mother face-to-face interaction: age and gender differences in coordination and the occurrence of miscoordination. *Child Dev, 60*(1), 85-92.

Tronick, E. Z., Als, H., Adamson, L., Wise, S., & Brazelton, B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *American Academy of Child Psychiatry*, *17*, 1-13.