DOI: 10.1111/infa.12556

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

RESEARCH ARTICLE

Automated measurement of infant and mother Duchenne facial expressions in the Face-to-Face/Still-Face

Yeojin Amy Ahn[1] | **Itir Önal Ertuğrul**[2] | **Sy-Miin Chow**[3] | **Jeffrey F. Cohn**[4] | **Daniel S. Messinger**[1,5,6]

[1]Department of Psychology, University of Miami, Coral Gables, Florida, USA

[2]Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

[3]Department of Human Development and Family Studies, Pennsylvania State University, State College, Pennsylvania, USA

[4]Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[5]Department of Electrical and Computer Engineering, University of Miami, Coral Gables, Florida, USA

[6]Departments of Pediatrics and Music Engineering, University of Miami, Coral Gables, Florida, USA

Correspondence
Daniel S. Messinger, Department of Psychology, University of Miami, 5665 Ponce de Leon Blvd., P.O. Box 248185, Coral Gables, FL 33124, USA.
Email: dmessinger@miami.edu

Funding information
National Institutes of Health, Grant/Award Numbers: GM105004, MH096951, U24 AA027684, UL1 TR002014-06, IGE 1806874, SES 1823633; National Science Foundation, Grant/Award Number: 1052736

Abstract
Although still-face effects are well-studied, little is known about the degree to which the Face-to-Face/Still-Face (FFSF) is associated with the production of intense affective displays. Duchenne smiling expresses more intense positive affect than non-Duchenne smiling, while Duchenne cry-faces express more intense negative affect than non-Duchenne cry-faces. Forty 4-month-old infants and their mothers completed the FFSF, and key affect-indexing facial Action Units (AUs) were coded by expert Facial Action Coding System coders for the first 30 s of each FFSF episode. Computer vision software, automated facial affect recognition (AFAR), identified AUs for the entire 2-min episodes. Expert coding and AFAR produced similar infant and mother Duchenne and non-Duchenne FFSF effects, highlighting the convergent validity of automated measurement. Substantive AFAR analyses indicated that both infant Duchenne and non-Duchenne smiling declined from the FF to the SF, but only Duchenne smiling increased from the SF to the RE. In similar fashion, the magnitude of mother Duchenne smiling changes over the FFSF were 2–4 times greater than non-Duchenne smiling changes. Duchenne expressions appear to be a sensitive index of intense infant and mother affective valence that

are accessible to automated measurement and may be a target for future FFSF research.

1 | INTRODUCTION

The Face-to-Face/Still-Face (FFSF) assesses infant responses to parent unresponsiveness, an age-appropriate stressor. Although still-face effects are well-studied, little is known about the degree to which the FFSF is associated with the production of intense affective displays, such as Duchenne smiles and Duchenne cry-faces that involve pronounced eye constriction. Duchenne smiles appear to express more intense positive affect than non-Duchenne smiles, while Duchenne cry-faces appear to express more intense negative affect than non-Duchenne cry-faces. The current study investigated whether levels of Duchenne and non-Duchenne smiling and cry-faces differ across FF, SF, and RE episodes, using computer vision-based automated measurement.

The Face-to-Face/Still-Face (FFSF; Tronick et al., 1978) has been used to examine how infants and parents communicate during naturalistic face-to-face interaction (Cohn & Tronick, 1987), how infants respond to parent emotional unavailability (Ekas et al., 2013), and how infant-parent dyads repair their interaction after the still-face (Feldman et al., 2010). In the FFSF, the parent is instructed to play with the infant normally (face-to-face) and to then stop responding to the infant and hold a neutral expression (i.e., "poker face") while maintaining eye contact (still-face), and to resume interaction with the infant (reunion). The FFSF is typically conducted in a laboratory context with the infant seated in an infant seat and the parent seated facing the infant. The prototypical still-face effect is characterized by reduced infant positive affect and gaze at the parent and increased negative affect during the still-face episode from the previous face-to-face episode (Mesman et al., 2009). The still-face effect tends to be followed by a partial carry-over effect into the reunion episode consisting of decreased levels of positive and increased levels of negative affect compared to the face-to-face.

The FFSF assesses infant response to the unresponsiveness, cessation of facial cues, and absence of contingent interaction in the still-face, and addresses whether infants return to positive interaction after the still-face, an index of self-regulation (Mesman et al., 2009; Tronick et al., 1978). The FFSF has been used to investigate developmental trajectories of infants' emotional regulation (Yato et al., 2008), episodic memory (Montirosso et al., 2013, 2014), physiological reactivity (Provenzi et al., 2017), genetic differences in infant self-regulation (Montirosso et al., 2016), self-regulation in infants of depressed and anxious mothers (Reck et al., 2013; Weinberg et al., 2006), and associations of early self-regulation with later externalizing and internalizing behaviors (Ekas et al., 2013; Moore et al., 2001). The still-face effect is robust to procedural variations (Mesman et al., 2009) and is associated with later attachment security (Ekas, et al., 2013).

FFSF research typically relies on manual coding of infants' positive (e.g., smiling, joy, positive vocalizations) and negative affect (e.g., frowning, crying, anger, sadness, fussing, protesting), as well as mothers' positive affect (e.g., smiling). The laborious quality of manual coding represents a limitation to sample sizes and the detail with which facial expressions are characterized (Cohn & Kanade, 2007; Messinger et al., 2012; Mitsven et al., 2020). Automated measurement approaches have the potential to address these limitations. The current study examined patterns of infant and mother Duchenne and non-Duchenne expressions in the FFSF and ascertained the convergent validity of expert coding and automated measurement of those patterns.

1.1 | Automated measurement

A growing body of research on emotional facial expressions has employed automated measurement using computer vision approaches, based on the Facial Action Coding System (FACS; Ekman et al., 2002) and its application to infants (BabyFACS; Oster, 2006). FACS is a comprehensive manual system for recording anatomically based appearance changes in the form of facial action units (AUs; Lucey et al., 2007), which has been used to intense positive affect and negative affect (Bolzani-Dinehart et al., 2005; Mattson, Cohn, et al., 2013; Messinger et al., 2009, 2012). Smiling (indexed by AU12) is the prototypical expression of positive emotion. Smiling combined with eye constriction (AU6; with associated raising of the cheeks), which is known as Duchenne smiling, indexes intense joy in adults (Ekman et al., 1990) and infants (Fox & Davidson, 1988). The cry-face (indexed by AU20), on the other hand, is the prototypical expression of negative emotion in infants (but not adults). Infant cry-faces involving stronger eye constriction, referred to as Duchenne cry-faces, are perceived to be more affectively negative than cry-faces with lower levels of eye constriction (Bolzani-Dinehart et al., 2005; Messinger, 2002; Oster, 2003).

A small body of research has examined infant Duchenne expression in the FFSF. Messinger et al. (2012) found that automated measurement of eye constriction (AU6; the Duchenne marker) indexed both the positive emotional intensity of smiling (AU12) and the negative emotional intensity of cry-faces (AU20) in twelve 6-month-olds. In the same sample, Mattson, Cohn, et al. (2013) showed that the proportion of smiling accompanied by eye constriction was higher during the positive-emotion eliciting face-to-face episode of the FFSF than during the still-face, and the proportion of cry-faces with eye constriction (stronger cry-faces) was higher during the negative-emotion eliciting still-face episode than during the face-to-face. These findings indicate that smiling with eye constriction (Duchenne smiling) indexes strong positive emotion in 6-month-old infants while cry-faces with eye constriction (Duchenne cry-faces) index strong negative emotion.

It remains unclear, however, whether infant Duchenne expressions are especially impacted by the FFSF procedure. One possibility, for example, is that the FFSF is associated with changes in levels of infant Duchenne expressions but has little or no bearing on non-Duchenne expressions. We are not aware of investigations of studies of mothers' Duchenne smiles in the FFSF. In adults, Duchenne smiles are perceived as intensified positive affective state and more genuine than non-Duchenne smiles, especially if they are naturally elicited, rather than posed (Gunnery & Ruben, 2016). It is unknown, however, whether mother Duchenne smiling is especially impacted by the FFSF in comparison with non-Duchenne smiling. To our knowledge, the current study is the first to document mother Duchenne smiling in the FFSF.

Automated detection of early facial expressions has yielded insights into the objective measurement of emotion expression during naturalistic interactions (Messinger et al., 2009, 2012). To date, however, the few applications to automated measurement to infant face-to-face interaction have used relatively small sample sizes. Moreover, no studies have reported convergent validity in which two measurements show the same changes in facial expressions in a specific standardized protocol. Assessing convergent validity of automated measurement of Duchenne and non-Duchenne expressions during the standardized FFSF protocol is a step toward reducing the human resource burden involved in making precise measurements of facial expressions in a large number of infants and mothers.

1.2 | Current study

The current study examined whether Duchenne and non-Duchenne expressions in 4-month-old infants are differentially distributed in the face-to-face, reunion, and still-face episodes of the FFSF.

We investigated whether expert manual coding and automated measurement (using computer vision approach) showed concordant changes in infant Duchenne and non-Duchenne smiling (AU12 with and without AU6, respectively) and in infant Duchenne and non-Duchenne cry-faces (AU4+AU20 with and without AU6, respectively) during the FFSF. Likewise, we examined whether expert coding and automated measurement of mother Duchenne and non-Duchenne smiling produce identical patterns of changes between the face-to-face and the reunion episodes. Finally, we assessed whether automated measurement of infant and mother Duchenne and non-Duchenne smiling–and infant Duchenne and non-Duchenne cry-faces–differed across the face-to-face, still-face, and reunion episodes. A pattern of results of particular interest would suggest that Duchenne smiling and Duchenne cry-faces are more sensitive indices of affective reactions to the FFSF than are non-Duchenne expressions.

2 | METHODS

2.1 | Participants

Forty 4-month-old (M_{age} = 4.07 months, SD = 0.31) full-term infants and their mothers (M_{age} = 29.38 years, SD = 5.15) participated in the study. Twenty-two infants (55%) were male, and 27 infants (68%) were Hispanic/Latinx. Infants' racial distribution composed of 28 White (70%), 7 Black/African American (17.5%), 4 multiracial (10%), and 1 unknown (2.5%). Twenty-three mothers (57.5%) were Hispanic/Latinx. Mothers' racial distribution consisted of 27 White (67.5%), 7 Black/African American (17.5%), 1 multiracial (2.5%), and 5 unknown (12.5%). Sample demographics are reported in Table 1. Data were collected between 2011 and 2014.

The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the Institutional Review Board at the University of Miami and conducted in accordance with the ethical standards of the American Psychological Association. Families were compensated for their participation.

2.2 | Procedure

Infants and their mothers completed the Face-to-Face/Still-Face (FFSF), which consisted of a 2-min face-to-face (FF), a 2-min still-face (SF), and a 2-min reunion (RE). Infants were placed in an elevated infant seat designed in-house, and mothers sat facing the infants. Mothers received instructions for each episode before the procedure ("This will be a 6 min activity, where you play with your baby for 2 min, put on a still-face and not interact with your baby for 2 min, then play with your baby again for 2 min"). They were instructed not to move out of the seat, to touch the infant's face, to lift infant up from the chair, or to shake infant from side to side or back and forth. They were informed that there will be audio prompting each episode.

The FF episode began with an audio recording, *Please play with your baby now*, followed by a 2-s tone. After 2 min, an audio recording, *Please stop interacting with your baby now*, and a 2-s tone prompted the SF episode. After 2 min, mothers heard another audio recording, *Please resume playing with your baby now*, followed by a 2-s tone, which indicated the beginning of RE episode. The interaction was recorded with one video camera directed at the infant's face and another at the mother's face and upper torso. Infant and mother videos were synchronized with a common time code.

TABLE 1 Demographics of infants and mothers.

	N (%)
Infant age (in months)	*M* = 4.07, SD = 0.31, Range = 3.4–4.7
Infant sex	
Male	22 (55%)
Female	18 (45%)
Infant ethnicity	
Hispanic/Latinx	27 (68%)
Other	13 (32%)
Infant race	
White	28 (70%)
Black/African-American	7 (17.5%)
Multiracial	4 (10%)
Unknown	1 (2.5%)
Mother age (in years)	*M* = 29.38, SD = 5.15, Range = 18–41
Mother ethnicity	
Hispanic/Latinx	23 (57.5%)
Other	17 (42.5%)
Mother race	
White	27 (67.5%)
Black/African-American	7 (17.5%)
Multiracial	1 (2.5%)
Unknown	5 (12.5%)
Mother education	
High school	7 (17.5%)
Some college	5 (12.5%)
2-year college	11 (27.5%)
4-year college	10 (25%)
Advanced/professional degree	6 (15%)
Unknown	1 (2.5%)

2.3 | Measures

2.3.1 | Manual coding

The Facial Action Coding System (FACS; Ekman et al., 2002) is an anatomically-based gold standard for measuring facial action units that has been adapted for use with infants (BabyFACS; Oster, 2006). Frame-by-frame anatomically based coding of mother and infant facial movement was conducted independently by one of three FACS-certified, Baby-FACS-trained coders. Smiling was indexed by the action of zygomaticus major (AU12), which pulls the lip corners laterally and upward; cry-faces were indexed by the combination of the action of risorius (AU20), which pulls the lip corners laterally, and the actions of depressor glabellae, depressor supercilia, and corrugator supercilli (AU4), which lowers the eyebrows. Eye constriction was indexed by the action of orbicularis

TABLE 2 Free-margin kappas (*S*-scores).

	AU4	AU6	AU12	AU20
Expert-expert infant	0.94	0.80	0.85	0.88
Expert-expert mother		0.82	0.92	
Expert-automated infant	0.88	0.75	0.77	0.81
Expert-automated mother		0.63	0.84	

Note: Overall free-margin kappas (*S*-scores) assessing reliability between expert FACS coders' manual coding (expert-expert) and between manual coding and automated measurement (expert-automated). Statistics were calculated over the first 30 s of FFSF episodes. $N = 40$.

oculi, pars orbitalis (AU6), which draws the cheeks and skin around the temples toward the eyes. Duchenne smiling involved the presence of both AU6 and AU12. Duchenne cry-faces involved the presence of AU4, AU6, and AU20. The individual AUs were coded as 0 (absent) or 1 (present). Duchenne smiling (AU6+AU12), non-Duchenne smiling (AU12 without AU6), Duchenne cry-faces (AU4+AU6+AU20), non-Duchenne cry-faces (AU4+AU20 without AU6) were coded as 1 (present) when all of their respective AUs were coded as 1.

Frame-by-frame expert coding of infant Duchenne and non-Duchenne smiling, and Duchenne and non-Duchenne cry-faces was obtained for the first 900 frames (30 s) of each episode (FF, SF, and RE). Expert coding of mother Duchenne and non-Duchenne smiling was obtained for the first 900 frames (30 s) of the FF and RE episodes. Expert coding of mother smiling was not obtained for the SF episode, based on the parental instructions to maintain a neutral facial expression during the SF. Sixteen infant videos (16 episodes) and 7 mother videos (7 episodes) were randomly selected to be coded by the other two FACS-certified, BabyFACS-trained coders. Intercoder reliability was examined using free-margin kappa (Brennan & Prediger, 1981) for each AU (AU4, AU6, AU12, and AU20). The mean free-margin kappa for infant AUs ($\kappa_{free} = 0.85$) were reported in Onal Ertuğrul et al. (2022). All free-margin kappas were above 0.80 (see first two rows of Table 2).

2.3.2 | Automated measurements

Infants' and mothers' synchronized video-recordings were submitted to the Automated Facial Affect Recognition (AFAR; Onal Ertuğrul, Cohn, et al., 2019, Onal Ertuğrul, Jeni, et al., 2019, 2022) computer vision software, an automated tool for detecting FACS AUs in the infant and mother faces. First, the Zface module (Jeni et al., 2017) of AFAR toolbox was used for automatic face tracking and registration. Zface used a fast-cascade regression framework to accomplish dense 3D registration of the face from the 2D video without person-specific training. The tracked facial images were normalized and centered in regard to rotation and scale to the inter-ocular distance (IOD) of 80 pixels, yielding 224 × 224 pixel images of faces with 80 pixels IOD. Mothers' facial images were then input to the Adult AFAR network, which is a convolutional neural network that was trained on adult faces (Onal Ertuğrul, Jeni, et al., 2019). Infants' facial images were input to the Infant AFAR network that was trained on infant faces (Onal Ertuğrul et al., 2022). The AFAR network output the probability of the facial AUs (i.e., AU4, AU6, AU12, AU20) for each video frame. A probability score of 0.5 or higher indicated the presence of an AU (recoded as 1). Using 25% of the frame-by-frame data (the first 30 s of each episode for each dyad), we obtained free-margin kappas above 0.63 for all reported infant and mother AUs between objective AFAR measurement and expert coding (see the last two rows of Table 2). These represent a moderate level of agreement corrected for chance. The kappa values of infant AUs were reported in Onal Ertuğrul et al. (2022).

In analyses, the proportions of expert-coded Duchenne and non-Duchenne smiling and Duchenne and non-Duchenne cry-faces in the first 30 s of each episode were compared to AFAR (automated) measurement. The proportions were computed as the number of frames of facial expression in the first 30 s of an episode divided by 900 frames, where facial expression was either Duchenne smiling (AU6+AU12), non-Duchenne smiling (AU12 without AU6), Duchenne cry-face (AU4+AU6+AU20), or non-Duchenne cry-face (AU4+AU20 without AU6).

For analyses of automated measurement data from the full 2-min episodes, the proportions were computed as the number of frames of facial expression in each episode divided by the total number of frames in each episode, where facial expression was either Duchenne smiling (AU6+AU12), non-Duchenne smiling (AU12 without AU6), Duchenne cry-face (AU4+AU6+AU20), or non-Duchenne cry-face (AU4+AU20 without AU6).

2.4 | Analytic approach

Analyses were conducted in IBM SPSS Statistics version 26.0.0 on a Windows computer. Using data from the first 30-s of each episode, we conducted a series of repeated measures analysis of variances (ANOVAs) to assess the validity of automated measurement. The patterns of changes in expert coding and automated measurement of infant Duchenne and non-Duchenne smiling, Duchenne and non-Duchenne cry-faces, and mother Duchenne and non-Duchenne smiling were compared. Following comparisons of expert and automated coding, we used repeated measures ANOVAs to assess the changes in the proportions of Duchenne and non-Duchenne smiling using the automated measurement in the full FFSF (i.e., 2-min FF, 2-min SF, and 2-min RE) and to compare those changes in Duchenne and non-Duchenne smiling across episodes. Identical ANOVA models were used for comparing infant Duchenne and non-Duchenne cry-faces and comparing mother Duchenne and non-Duchenne smiling.

3 | RESULTS

3.1 | Convergent validity of automated measurement

Missing data were comparable for expert coding and automated measurement. Data for infants included the first 30 s of the FF, SF, and RE episodes while data for mothers included only the FF and RE. Experts used an un-codable category for 30,523 frames (28.3%) of infant videos and 2699 frames (3.7%) of mother videos, due to occlusion (e.g., the infant or mother's hands blocked a view of their faces). In the same reliability sample, the automated Zface approach did not identify the face (and hence could not identify facial actions) in a total of 18,283 frames (16.9%) of infant videos and 1099 frames (1.5%) of mother videos. See Table S1 in Supplementary Materials S1 for missing data information for the full FFSF.

Table 3 includes descriptive statistics of the proportions of infant Duchenne and non-Duchenne smiling and cry-faces and mother Duchenne and non-Duchenne smiling, obtained by expert coding and automated measurement during the first 30 s of each episode of the FFSF. To assess the convergent validity of automated measurement, we conducted a series of 3 (Episode: FF, SF, RE) × 2 (Coding: Expert, Automated) mixed-model ANOVAs. Expert coding and automated measurement were compared in the changes in mean proportions of infant Duchenne and non-Duchenne smiling and Duchenne and non-Duchenne cry-faces across episodes (the first 30 s). These were followed by pairwise comparisons between episodes. The Least Significant Difference comparisons are reported

TABLE 3 Proportions of infant and mother Duchenne and non-Duchenne expressions.

	FF		SF		RE	
	Expert	Automated	Expert	Automated	Expert	Automated
Infant; mean (SD)						
Duchenne smiling	0.18 (0.18)	0.22 (0.20)	0.07 (0.08)	0.09 (0.09)	0.14 (0.15)	0.20 (0.19)
Non-Duchenne smiling	0.06 (0.07)	0.07 (0.08)	0.03 (0.06)	0.03 (0.06)	0.05 (0.09)	0.05 (0.06)
Duchenne cry-face	0.01 (0.02)	0.01 (0.04)	0.02 (0.08)	0.03 (0.11)	0.08 (0.18)	0.09 (0.20)
Non-Duchenne cry-face	0.00 (0.00)	0.03 (0.06)	0.01 (0.03)	0.06 (0.13)	0.005 (0.01)	0.13 (0.21)
Mother; Mean (SD)						
Duchenne smiling	0.63 (0.23)	0.67 (0.22)	-	0.11 (0.16)	0.47 (0.29)	0.51 (0.26)
Non-Duchenne smiling	0.26 (0.20)	0.22 (0.18)	-	0.18 (0.21)	0.27 (0.23)	0.27 (0.19)

Note: Means and (standard deviations) of the proportions of infant Duchenne and non-Duchenne smiling and cry-faces and mother Duchenne and non-Duchenne smiling during the first 30 s of each episode of the FFSF. $N = 40$.

in Supplementary Materials S1 (Tables S2 and S3). Results were identical when Bonferroni corrections were applied.

3.1.1 | Infant smiling

There were Episode effects for infant Duchenne, F (1.73, 67.63) = 12.06, $p < 0.001$, $\eta_p^2 = 0.24$, and non-Duchenne smiling, F (2, 78) = 5.94, $p = 0.004$, $\eta_p^2 = 0.13$ (see Table 4). For Duchenne smiling, there was a Coding effect that indicated a greater proportion of automated measure of Duchenne smiling than expert-coded Duchenne smiling, F (1, 39) = 13.22, $p = 0.001$, $\eta_p^2 = 0.25$. There was no Coding effect for non-Duchenne smiling, F (1, 39) = 0.001, $p = 0.98$, $\eta_p^2 < 0.001$. There were no interactions between Episode and Coding for either Duchenne smiling, F (2, 78) = 1.51, $p = 0.23$, $\eta_p^2 = 0.04$, or non-Duchenne smiling, F (2, 78) = 0.09, $p = 0.92$, $\eta_p^2 = 0.002$.

Follow-up Least Significant Difference post-hoc comparisons indicated that expert coding and automated measurement both showed declines in Duchenne and non-Duchenne smiling from the FF to the SF (ps < 0.05). Both expert coding and automated measurement indicated a rise in Duchenne smiling from the SF to the RE (ps < 0.01). Neither expert coding nor automated measurement indicated mean difference in the proportions of non-Duchenne smiling between the SF and RE episodes (see Table S2). That is, expert coding and automated measurement yielded identical significant infant still-face effects involving Duchenne and non-Duchenne smiling (see Figure 1).

3.1.2 | Infant cry-faces

There was an Episode effect for infant Duchenne cry-faces, F (1.41, 54.96) = 6.34, $p = 0.01$, $\eta_p^2 = 0.14$, and non-Duchenne cry-faces, F (1.81, 70.65) = 5.59, $p = 0.01$, $\eta_p^2 = 0.13$ (Table 4). There was no Coding effect for Duchenne cry-faces, F (1, 39) = 3.32, $p = 0.08$, $\eta_p^2 = 0.08$. However, there was a Coding effect for non-Duchenne cry-faces, F (1, 39) = 19.84, $p < 0.001$, $\eta_p^2 = 0.34$, such that there was a greater proportion of automated measure of non-Duchenne cry-faces than expert-coded non-Duchenne cry-faces. There was no interaction between Episode and Coding for infant Duchenne cry-faces, F (1.70, 66.34) = 0.17, $p = 0.81$, $\eta_p^2 = 0.004$. However, there was a significant interaction between Episode and Coding for non-Duchenne cry-faces, F (1.73, 67.53) = 6.07, $p = 0.01$, $\eta_p^2 = 0.13$.

For both expert coding and automated measurement, post-hoc pairwise comparisons indicated that the proportion of Duchenne cry-faces did not differ between the FF and the SF ($p = 0.27$) but

TABLE 4 Analysis of variance (ANOVA) comparisons of expert coding and automated measurement over the Face-to-Face/Still-Face (FFSF).

Measures	Test of sphericity	Adjustment method and ε level	Effect	F	df1	df2	p	η_p^2	Power (1-β error prob.)
Infant measures									
Duchenne smiling	χ^2 (2) = 8.42, p = 0.01	Huynh-Feldt correction (ε = 0.87)	Episode	12.06	1.73	67.63	<0.001	0.24	1.00
			Coding	13.22	1	39	0.001	0.25	1.00
	χ^2 (2) = 1.01, p = 0.61	Sphericity assumed (ε = 0.97)	Episode × coding	1.51	2	78	0.23	0.04	0.79
Non-Duchenne smiling	χ^2 (2) = 3.23, p = 0.20	Sphericity assumed (ε = 0.92)	Episode	5.94	2	78	0.004	0.13	1.00
			Coding	0.001	1	39	0.98	<0.001	0.05
	χ^2 (2) = 0.94, p = 0.63	Sphericity assumed (ε = 0.98)	Episode × coding	0.09	2	78	0.92	0.002	0.09
Duchenne cry-face	χ^2 (2) = 20.65, p < 0.001	Greenhouse-Geisser correction (ε = 0.70)	Episode	6.34	1.41	54.96	0.01	0.14	1.00
			Coding	3.32	1	39	0.08	0.08	0.99
	χ^2 (2) = 9.46, p = 0.01	Huynh-Feldt correction (ε = 0.85)	Episode × coding	0.17	1.70	66.34	0.81	0.004	0.12
Non-Duchenne cry-face	χ^2 (2) = 6.22, p = 0.04	Huynh-Feldt correction (ε = 0.91)	Episode	5.59	1.81	70.65	0.01	0.13	1.00
			Coding	19.84	1	39	<0.001	0.34	1.00
	χ^2 (2) = 8.50, p = 0.01	Huynh-Feldt correction (ε = 0.87)	Episode × coding	6.07	1.73	67.53	0.01	0.13	1.00
Mother measures									
Duchenne smiling			Episode	22.47	1	39	<0.001	0.37	1.00
			Coding	2.58	1	39	0.12	0.06	0.88
			Episode × coding	0.02	1	39	0.88	0.001	0.07

(Continues)

TABLE 4 (Continued)

Measures	Test of sphericity	Adjustment method and ε level	Effect	F	df1	df2	p	η_p^2	Power (1-β error prob.)
Non-Duchenne smiling			Episode	1.70	1	39	0.20	0.04	0.71
			Coding	0.46	1	39	0.50	0.01	0.24
			Episode × coding	1.95	1	39	0.17	0.05	0.81

Note: Analysis of variance (ANOVA) comparisons of the proportions of infant and mother facial expressions during the first 30 s of each episode of the FFSF. Episode: main effect of episode (FF, SF, RE for infants; FF, RE for mothers); Coding: main effect of coding (expert and automated); Episode × Coding: interaction effect of Episode by Coding. The assumption of sphericity was violated for the main effect of episode on infant Duchenne smiling, Duchenne and non-Duchenne cry-faces, as well as the interaction effect of episode by coding on infant Duchenne and non-Duchenne cry-faces. The degrees of freedom were corrected using a Huynh-Feldt correction if ε is greater than 0.75 and a Greenhouse-Geisser correction if ε is less than 0.75 (Field, 2013). The Episode effect result for infant Duchenne cry-faces is not altered when its degrees of freedom are corrected using a Huynh-Feldt correction. Post-hoc power analysis was performed using G*Power 3.1. Linear mixed-effects analyses yielded conclusions identical to those based on the current results. η_p^2 = Partial eta squared.

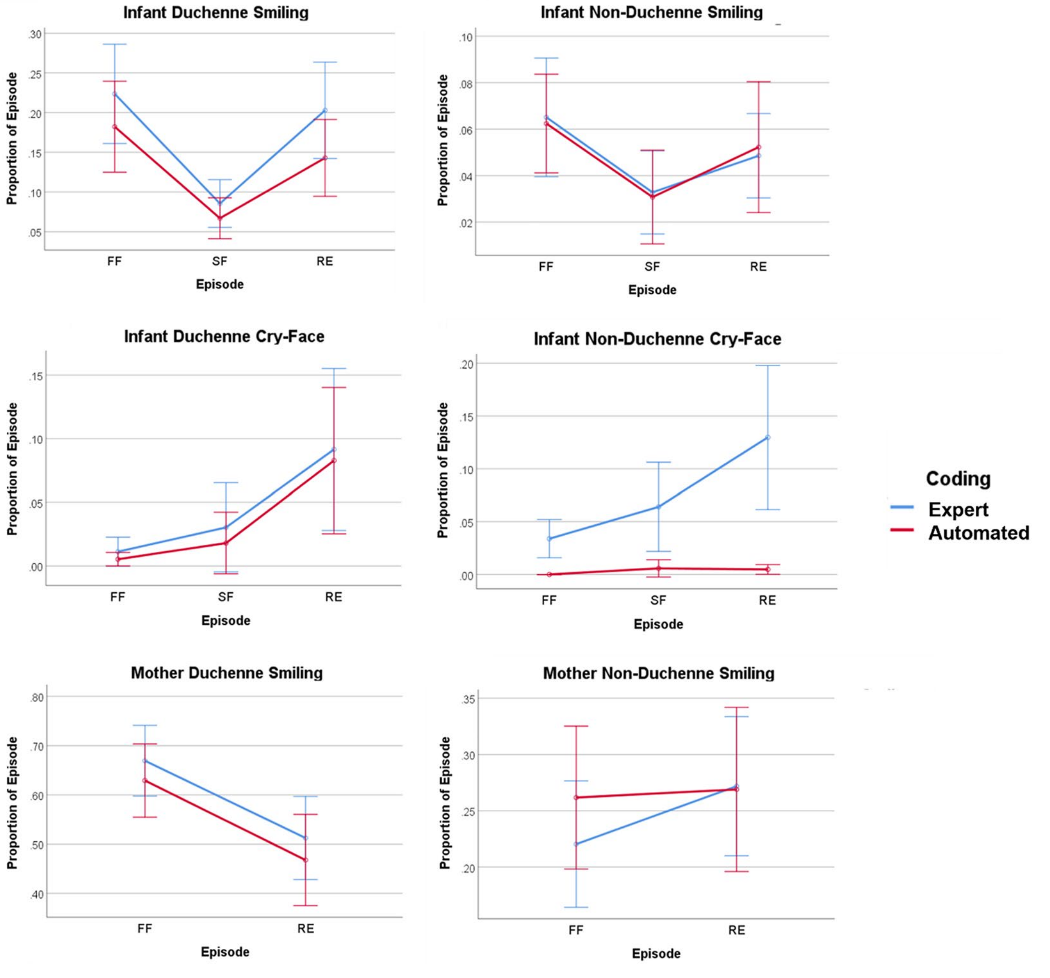

FIGURE 1 Expert coding and automated measurement of infant and mother Duchenne and non-Duchenne expressions in the first 30 s of each episode. Line plots contain estimated marginal means of the proportions of infant Duchenne and non-Duchenne smiling (first row), infant Duchenne and non-Duchenne cry-faces (second row), and mother Duchenne and non-Duchenne smiling (third row). The first column shows the Duchenne expressions, and the second column shows the non-Duchenne expressions. The *x*-axes indicate the Face-to-Face/Still-Face (FFSF) episodes (FF, SF, and RE for infants; FF and RE for mothers). The *y*-axes indicate mean proportions. The blue line represents automated facial affect recognition (AFAR) coding, and the red line represents expert FACS coding. Error bars indicate 95% confidence intervals.

increased from the SF to the RE ($p = 0.02$). Non-Duchenne cry-faces did not differ between the FF and the SF for either expert coding or automated measurement (*p*s > 0.05). Automated measurement, but not expert coding, yielded a significant rise in non-Duchenne cry-faces from the SF to the RE ($p = 0.04$). Post-hoc pairwise comparisons are reported in Table S2.

3.1.3 | Mother smiling

For mothers, 2 (Episode: FF, RE) × 2 (Coding: Expert, Automated) ANOVAs were used to compare the changes from the FF to RE in expert coding and automated measurement of Duchenne and non-Duchenne smiling. There was an Episode effect for mother Duchenne smiling, F (1, 39) = 22.47, $p < 0.001$, $\eta_p^2 = 0.37$, but not for non-Duchenne smiling, F (1, 39) = 1.70, $p = 0.20$, $\eta_p^2 = 0.04$. Specifically, for both expert coding and automated measurement, there was a decline in mother Duchenne smiling, but not in non-Duchenne smiling, from the FF to the RE (Figure 1). There was no Coding effect for Duchenne, F (1, 39) = 2.58, $p = 0.12$, $\eta_p^2 = 0.06$, and non-Duchenne smiling, F (1, 39) = 0.46, $p = 0.50$, $\eta_p^2 = 0.01$. There were no interactions between Episode and Coding for Duchenne, F (1, 39) = 0.02, $p = 0.88$, $\eta_p^2 = 0.001$, and non-Duchenne smiling, F (1, 39) = 1.95, $p = 0.17$, $\eta_p^2 = 0.005$.

3.1.4 | Associations between infant and mother smiling

We examined associations between levels of infant and mother Duchenne and non-Duchenne smiling in the first 30 s of the FF and RE episodes to ascertain whether expert and automated coding produced similar associations. Levels of infant Duchenne and non-Duchenne smiling were, respectively, not associated with mother Duchenne and non-Duchenne smiling in the first 30 s of the FF and RE episodes. The absence of significant associations was observed for both automated measurement and expert coding (automated Duchenne smiling: r [78] = 0.09, $p = 0.43$; expert-coded Duchenne smiling: r [78] = 0.13, $p = 0.24$; automated non-Duchenne smiling: r [78] = −0.16, $p = 0.17$; expert-coded non-Duchenne smiling: r [78] = −0.11, $p = 0.31$).

3.2 | Full automated measurement of FFSF effects

Table 5 includes means and standard deviations of the proportions of infant smiling and cry-faces, as well as mother smiling, during each of the *full 2-min episodes* of the FFSF, which were obtained via automated measurement. A series of 3 (Episode: FF, SF, RE) × 2 (Type: Duchenne, non-Duchenne) mixed-model ANOVAs were conducted.

3.2.1 | Infant smiling

There was an Episode effect for infant smiling, F (2, 78) = 22.46, $p < 0.001$, $\eta_p^2 = 0.37$, which indicated that there was an overall change in total smiling (Duchenne and non-Duchenne) across episodes. There was a Type effect for infant smiling, F (1, 39) = 44.43, $p < 0.001$, $\eta_p^2 = 0.53$, which indicated a greater proportion of Duchenne smiling than non-Duchenne smiling across episodes. Moreover, there was a significant interaction between Episode and Type, F (2, 78) = 11.19, $p < 0.001$, $\eta_p^2 = 0.22$, which indicated that the pattern of changes across the FFSF episodes differed between Duchenne and non-Duchenne smiling (Figure 2).

Pairwise comparisons indicated that total smiling declined from the FF to the SF ($p < 0.001$) and increased from the SF to the RE ($p < 0.001$). Both infant Duchenne and non-Duchenne smiling declined from the FF to the SF (ps < 0.001), but only Duchenne smiling rose significantly from the SF to the RE ($p < 0.001$). Post-hoc pairwise comparisons are reported in Table S3.

3.2.2 | Infant cry-faces

There was an Episode effect for infant cry-faces (Duchenne and non-Duchenne), F (2, 78) = 6.15, p = 0.003, η_p^2 = 0.14. There was no Type effect for cry-faces, which indicated that the proportions of infant Duchenne and non-Duchenne cry-faces did not differ across episodes, F (1, 39) = 1.56, p = 0.22, η_p^2 = 0.04. There was a significant interaction between Episode and Type for infant cry-faces,

TABLE 5 Analysis of variance (ANOVA) and descriptive statistics of automated facial affect recognition (AFAR)-identified facial expressions over the entire Face-to-Face/Still-Face (FFSF).

Measures	Effect		*M*	SD	*F*	df1	df2	*p*	η_p^2	Power (1-β error prob.)
Infant smiling	Episode	FF	0.12	0.13	22.46	2	78	<0.001	0.37	1.00
		SF	0.06	0.09						
		RE	0.10	0.12						
	Type	Duchenne	0.15	0.14	44.43	1	39	<0.001	0.53	1.00
		Non-Duchenne	0.04	0.05						
	Episode × type	FF Duchenne	0.20	0.15	11.19	2	78	<0.001	0.22	1.00
		FF non-Duchenne	0.05	0.06						
		SF Duchenne	0.09	0.11						
		SF non-Duchenne	0.02	0.04						
		RE Duchenne	0.17	0.14						
		RE non-Duchenne	0.03	0.04						
Infant cry-faces	Episode	FF	0.02	0.05	6.15	2	78	0.003	0.14	1.00
		SF	0.05	0.13						
		RE	0.06	0.13						
	Type	Duchenne	0.06	0.14	1.56	1	39	0.219	0.04	0.80
		Non-Duchenne	0.03	0.07						
	Episode × type	FF Duchenne	0.01	0.03	5.00	2	78	0.009	0.11	1.00
		FF non-Duchenne	0.03	0.06						
		SF Duchenne	0.07	0.17						
		SF non-Duchenne	0.04	0.07						
		RE Duchenne	0.10	0.17						
		RE non-Duchenne	0.03	0.06						
Mother smiling	Episode[a]	FF	0.42	0.26	167.10	1.44	56.28	<0.001	0.81	1.00
		SF	0.10	0.15						
		RE	0.36	0.22						
	Type	Duchenne	0.36	0.30	12.20	1	39	0.001	0.24	1.00
		Non-Duchenne	0.22	0.18						

(Continues)

TABLE 5 (Continued)

Measures	Effect		*M*	SD	*F*	df1	df2	*p*	η_p^2	Power (1-β error prob.)
	Episode × type	FF Duchenne	0.58	0.23	31.36	2	78	<0.001	0.45	1.00
		FF non-Duchenne	0.25	0.17						
		SF Duchenne	0.06	0.09						
		SF non-Duchenne	0.14	0.18						
		RE Duchenne	0.45	0.24						
		RE non-Duchenne	0.28	0.16						

Note: Statistics were computed using the proportions of infant smiling and cry-faces and mother smiling during the full 2-min episodes of the FFSF. Raw means and standard deviations are reported. Episode: main effect of episode (FF vs. SF vs. RE); Type: main effect of type of expression (Duchenne vs. Non-Duchenne expressions); Episode × Type: interaction effect of episode by type of expression. Post-hoc power analysis was performed using G*Power 3.1. η_p^2 = Partial eta squared.

[a]As the assumption of sphericity was violated for the main effect of episode on mother smiling (Mauchly's $W = 0.61$, $p < 0.001$), degrees of freedom were corrected using a Greenhouse-Geisser correction ($\varepsilon = 0.72$). Assumption of sphericity was met for the main effects and interaction effect on infant smiling and cry-faces, and for the interaction effect on mother smiling (Mauchly's *W*s > 0.90, *p*s > 0.05). Linear mixed-effects analyses yielded conclusions identical to those based on the current results.

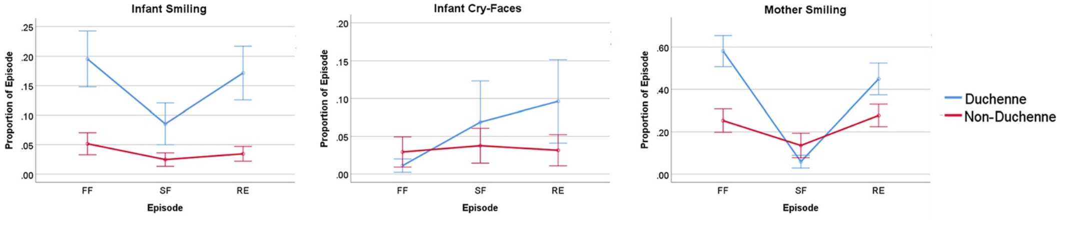

FIGURE 2 Full-episode automated measurement over the full Face-to-Face/Still-Face (FFSF). Estimated marginal means of the proportions of automated facial affect recognition (AFAR)-identified infant Duchenne and non-Duchenne smiling (first plot), infant Duchenne and non-Duchenne cry-faces (second plot), and mother Duchenne and non-Duchenne smiling (third plot) in each full 2-min episode of the FFSF. Time in infant smiling, infant cry-faces, and mother smiling as a proportion of time in each episode of the FFSF. The blue line represents Duchenne expressions, and the red line represents non-Duchenne expressions. The first plot demonstrates significant declines in both infant Duchenne and non-Duchenne smiling from the FF to the SF, and a rise in infant Duchenne smiling from the SF to the RE. The second plot illustrates a significant rise in infant Duchenne cry-faces from the FF to the SF and no change from the SF to the RE. There was no difference in infant non-Duchenne cry-faces between the episodes. The third plot shows significant declines in mother Duchenne and non-Duchenne smiling from the FF to the SF, and a rise from the SF to the RE. Error bars indicate 95% confidence intervals. Note that the *y*-axis scales are different for infant smiling, infant cry-faces, and mother smiling.

F (2, 78) = 5.00, p = 0.01, η_p^2 = 0.11, which indicated that the pattern of changes across the FFSF episodes differed between Duchenne and non-Duchenne cry-faces (Figure 2).

Pairwise comparisons indicated that total cry-faces (Duchenne and non-Duchenne) increased from the FF to the SF ($p = 0.02$) but did not change significantly between the SF and RE ($p = 0.40$), which appears to be driven by Duchenne cry-faces, as Duchenne cry-faces increased from the FF to the

SF ($p = 0.03$) and did not change between the SF and the RE ($p = 0.27$; Table S3). Non-Duchenne cry-faces did not change significantly between the FF and the SF ($p = 0.28$), or between the SF and the RE ($p = 0.19$).

3.2.3 | Mother smiling

There was an Episode effect for mother smiling (Duchenne and non-Duchenne), F (1.44, 56.28) = 167.10, $p < 0.001$, $\eta_p^2 = 0.81$. Similar to infant smiling, there was a Type effect, F (1, 39) = 12.20, $p = 0.001$, $\eta_p^2 = 0.24$, which indicated a greater proportion of Duchenne smiling than non-Duchenne smiling across episodes. Moreover, there was a significant interaction between Episode and Type, F (2, 78) = 31.36, $p < 0.001$, $\eta_p^2 = 0.45$, which indicated that the pattern of changes across the episodes differed between Duchenne smiling and non-Duchenne smiling (Figure 2).

Pairwise comparisons indicated that total smiling (Duchenne and non-Duchenne) declined from the FF to the SF ($p < 0.001$), then rose from the SF to the RE ($p < 0.001$). Both Duchenne and non-Duchenne smiling significantly declined from the FF to the SF ($p < 0.001$ and $p = 0.002$, respectively) and rose from the SF to the RE (ps < 0.001). However, the mean differences between the episodes were 2–4 times greater for Duchenne smiling than for non-Duchenne smiling (see Table S3).

4 | DISCUSSION

The FFSF is widely used to examine infant affective responses to parental unresponsiveness. However, we have little understanding of the role of Duchenne expressions, which index strong affective states, in infants and parents during the FFSF. This may be due to the resources required for manual coding of facial AUs such as AU6 (cheek raising or eye constriction) that characterizes Duchenne expressions. The current study examined infant and mother Duchenne and non-Duchenne expressions in the FFSF, using both manual expert coding and an automated measurement procedure based on machine learning detection to detect AUs. Expert coding and automated measurement of FACS AUs yielded similar patterns of change in infant and mother Duchenne and non-Duchenne expressions across the FFSF. These results highlight the convergent validity of automated measurement of facial expressions, particularly smiling, in the FFSF.

Full-episode AFAR data indicated that both infants and mothers demonstrated more Duchenne smiling during the interactive FF and RE episodes than the SF episode. Infants demonstrated more Duchenne cry-faces during the SF than the interactive FF and RE episodes. However, these patterns were not evident for infant non-Duchenne cry-faces. These findings suggest that Duchenne expressions may be an especially sensitive index of infant and mother affective valence during the FFSF. It is possible, then, that standard finding of reductions of infant smiling and increases in cry-faces during the FFSF are specifically dependent on changes in the Duchenne forms of these expressions.

4.1 | Expert coding and automated measurement

We ascertained the convergent validity of infant and mother Duchenne and non-Duchenne smiling—and infant Duchenne and non-Duchenne cry-faces—by testing for differences between expert and automated (AFAR) levels of these expressions (main effects of measurement approach). There were no differences between expert coding and AFAR in the overall proportion of infant non-Duchenne smiling and Duchenne cry-faces, which suggests convergent validity for these expressions. However, AFAR detected more infant Duchenne smiling and non-Duchenne cry-faces than did expert coders. In fact, expert coding indicated almost no non-Duchenne cry-faces in the FFSF. It is noteworthy, however, that there was substantial agreement between expert coding and AFAR at an individual AU level (Kappas above 0.75). One possibility is that AFAR over-detects AU6 during smiles and under-detects the absence of AU6 during cry-faces. This suggests AFAR does not reliably detect all facial expressions formed by different AU combinations. It may be that data augmentation of training data for machine learning is necessary to more accurately measure rarely occurring combinations of AUs, such as those involved in the non-Duchenne cry-face. AU6, termed cheek raiser in FACS, constricts the eyes, which may be a relatively subtle change in infants who have relatively high levels of subcutaneous fat. Another possibility is that AFAR is more accurate at detecting these changes than human coders. Further investigation employing automated measurement appropriately trained on combinations of AUs is warranted.

We also examined patterns of change in infant and mother Duchenne and non-Duchenne expressions across FFSF episodes (the interaction between measurement approach and episode). The convergent validity of AFAR automated measurement was supported by non-significant interaction effects for infant Duchenne and non-Duchenne smiling, and Duchenne cry-faces. This suggests that expert coding and automated measurement yielded substantively identical FFSF effects for these infant expressions. However, a significant interaction effect for infant non-Duchenne cry-faces indicated that while automated measurement indicated an increase in infant non-Duchenne cry-faces from the SF to the RE, expert coding indicated no change between the two episodes. A conservative interpretation suggests caution in interpreting FFSF effects based on automated measurement of non-Duchenne cry-faces.

For mothers, there were no overall differences in the levels of Duchenne and non-Duchenne smiling between expert coding and AFAR. In addition, interaction effects were not significant. Reliability between expert and AFAR coding for mother facial expressions were somewhat variable ($\kappa = 0.63$ for AU6 and $\kappa = 0.84$ for AU12). Nevertheless, results from both expert coding and AFAR indicated declines in Duchenne smiling in the RE compared to the FF for expert coding and automated measurement. For both expert coding and automated measurement, mother non-Duchenne smiling did not differ between the FF and the RE. This suggests a concordance between expert coding and automated measurement for mother smiling in the FFSF. Levels of infant and mother Duchenne and non-Duchenne smiling were not associated during the first 30 s of the FFSF. Although this finding may reflect the brief sample window examined, it was consistent between experts and AFAR. It appears that expert coding and automated measurement may produce broadly comparable results when applied to one of the most common interactive protocols used in early infancy.

4.2 | Automated Duchenne and non-Duchenne expression

4.2.1 | Infant smiling

Here we report on substantive results from automated facial expression detection (AFAR) in the full FFSF episodes. Infant total smiling (Duchenne and non-Duchenne) declined from the FF to the SF. This finding is consistent with reports of declines in overall infant smiling from the FF to the SF (Ekas et al., 2013; Mattson, Cohn, et al., 2013, Mattson, Ekas, et al., 2013) and with more general reports of declines in of positive affect indicators from the FF to the SF (Mesman et al., 2009; Stockdale et al., 2020; Yaari et al., 2018). We found that both infant Duchenne smiling and non-Duchenne smiling declined from the FF to the SF, which is consistent with Mattson, Cohn, et al. (2013) report on 12 infants. However, Mattson et al. reported on Duchenne smiling as a proportion of overall smiling, while we examined Duchenne smiling as a proportion of the overall episode. We believe the overall proportion to be a relatively convenient, well-validated metric of infant positive affect, which is well suited to the exploration of individual differences in the FFSF.

Automated measurement of the FF, SF, and RE episodes indicated accentuated patterns of changes across episodes for infant Duchenne expressions. Only Duchenne smiling (not non-Duchenne smiling) increased from the SF to the RE. Thus, infants demonstrated more Duchenne smiling in the interactive FF and RE episodes than in the SF. In other words, when infants interacted with their parents, infant smiles were more emotionally positive than when reacting to a still-faced parent. In addition, a significant interaction effect indicated that there was a steeper decline in Duchenne smiling than in non-Duchenne smiling from the FF to the SF and a steeper increase from the SF to the RE. The results suggest that infant Duchenne smiling is an especially sensitive metric of response to the cessation and resumption of playful interaction with the parent. To index infant positive affect in the FFSF future studies might efficiently focus on Duchenne smiling—rather than all infant smiling.

4.2.2 | Infant cry-faces

The current study extended Mattson, Cohn, et al. (2013) investigation of infant Duchenne and non-Duchenne cry-faces in the FF and SF to the RE episode. Automated measurement indicated a rise in infant Duchenne cry-faces, but not non-Duchenne cry-faces, from the FF to the SF. This is consistent with Mattson, Cohn, et al. (2013) reports of greater levels of the Duchenne marker in cry-faces in the negative affect-eliciting SF episode than the positive affect-eliciting play (FF) episode, as well as with more general reports of increases in infant cry-faces from the FF to the SF (Ekas et al., 2013; Mattson, Cohn, et al., 2013, Mattson, Ekas, et al., 2013). Infant Duchenne cry-faces persisted into the RE episode of the current FFSF. In the RE episode, infants typically show a partial reduction in affectively negative behaviors from the previous SF episode (Mesman et al., 2009). These results suggest that the infant Duchenne cry-face is a particularly sensitive metric of the impact of the cessation of playful interaction with the parent on subsequent interaction. Future research might focus on Duchenne cry-faces—rather than all infant cry-faces—to index infant negative facial affect in the FFSF.

4.2.3 | Mother smiling

Validating adherence to instructions to hold a still-face, we found both greater mother Duchenne and non-Duchenne smiling in the FF and the RE than in the SF. Following up on a significant interaction effect, mothers showed proportionally more Duchenne smiling than non-Duchenne smiling in the FF and the RE than the SF. Naturally-elicited Duchenne smiling appears to index more intense positive affect than non-Duchenne smiling in adults (Gunnery & Ruben, 2016). The current findings of accentuated levels of mother Duchenne smiling during the interactive episodes of the FFSF suggest that Duchenne smiling may function as a primary expression of positive affect for mothers in infant-mother face-to-face interaction (Messinger et al., 2009).

4.3 | Limitations and future directions

The current study compared expert and automated approaches to investigate infant and mother Duchenne and non-Duchenne expressions in a standard interaction protocol. Accentuated patterns of changes in Duchenne expressions over the FFSF suggest that still-face effects may be driven by Duchenne expressions that index intense positive and negative affect for both infants and mothers in infant-mother face-to-face affective communication. While the Duchenne marker was coded as present or absent, it is also a continuous action produced by the muscle encircling the eye (Messinger et al., 2009, 2012). Future investigations employing continuous measurements of the *intensity* of the Duchenne marker and other facial actions may further our understanding of facial expressions during infant-mother face-to-face interactions and dynamic changes within episodes of the FFSF (Ekas et al., 2013).

We used experts who were trained in and employed the comprehensive, anatomically-based FACS to code facial actions (Ekman et al., 2002). Based on that standard, we conducted reliable, automated measurement of facial expressions in the largest FFSF sample to date. Levels of missing data in which facial actions could not be coded were comparable for the expert and automated approaches. For both approaches, missing data levels were higher for infants who exhibit high levels of head and hand movement—than for mothers. Nevertheless, there are limitations to the current study's automated measurement approach.

Although automated measurement reliably measured facial expressions, the machine learning algorithms required expert coders' manual measurement for training, which may replicate subjectivity and biases involved in manual coding. Moreover, limitations in manually coded training data may constrain the performance and function of automated measurement when applied to different samples in different contexts. For example, mothers, but not fathers or other caregivers, were participants in the current study. Likewise, the FFSF physically constrains the infant and mother, and toys were not permitted. Thus, the current results may differ from those in which infants and parents play on the floor, or play with objects. The development of automated measurement methods trained with diverse samples in diverse settings is imperative for improving the objectivity and generalizability of automated measurement methods. Potential applications include infant temperament assessments and clinical protocols for assessing behavioral indicators of disorders such as autism spectrum disorder (Ahn et al., 2023; Carpenter et al., 2021).

In sum, the current study united research on intense affective valence communicated by Duchenne expressions with research on the widely used FFSF assessment of socioemotional functioning competence (Granat et al., 2017; Mitsven et al., 2020). Results highlight the potential of computer vision approaches to enable objective understanding of early interactive facial expressions. They suggest that affective changes produced by the FFSF may be uniquely reflected in infant and mother Duchenne expressions, and that these expressions are indices of intense positive and negative emotion.

ACKNOWLEDGMENTS

The project described was supported in part by National Institutes of Health grants MH096951, GM105004, U24 AA027684, UL1 TR002014-06, IGE 1806874, SES 1823633, and National Science Foundation grant 1052736. We thank the families who participated. The authors declare no conflicts of interest with regard to the funding source for this study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Databrary at https://nyu.databrary.org/volume/1474.

ORCID

Yeojin Amy Ahn https://orcid.org/0000-0003-3504-6909

REFERENCES

Ahn, Y. A., Moffitt, J. M., Tao, Y., Custode, S., Parlade, M., Beaumont, A., Cardona, S., Hale, M., Durocher, J., Alessandri, M., Shyu, M., Perry, L. K., & Messinger, D. S. (2023). Objective measurement of social gaze and smile behaviors in children with suspected autism spectrum disorder during administration of the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*. 2nd ed. https://doi.org/10.1007/s10803-023-05990-z

Bolzani-Dinehart, L., Messinger, D., Acosta, S., Cassel, T., Ambadar, Z., & Cohn, J. (2005). Adult perceptions of positive and negative infant emotional expressions. *Infancy*, *8*(3), 279–303. https://doi.org/10.1207/s15327078in0803_5

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, *41*(3), 687–699. https://doi.org/10.1177/001316448104100307

Carpenter, K. L. H., Hahemi, J., Campbell, K., Lippmann, S. J., Baker, J. P., Egger, H. L., Espinosa, S., Vermeer, S., Sapiro, G., & Dawson, G. (2021). Digital behavioral phenotyping detects atypical pattern of facial expression in toddlers with autism. *Autism Research*, *14*(3), 488–499. https://doi.org/10.1002/aur.2391

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. 22–238). Oxford University Press.

Cohn, J. F., & Tronick, E. Z. (1987). Mother–infant face-to-face interaction: The sequence of dyadic states at 3, 6, and 9 months. *Developmental Psychology*, *23*(1), 68–77. https://doi.org/10.1037/0012-1649.23.1.68

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. https://doi.org/10.1037/a0029330

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, *58*(2), 342–353. https://doi.org/10.1037/0022-3514.58.2.342

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system investigator's guide*. A Human Face.

Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, *13*(2), 271–278. https://doi.org/10.1111/j.1467-7687.2009.00890.x

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.

Fox, N. A., & Davidson, R. J. (1988). Patterns of brain electrical activity during facial signs of emotion in 10-month-old infants. *Developmental Psychology*, *24*(2), 230–236. https://doi.org/10.1037/0012-1649.24.2.230

Granat, A., Gadassi, R., Gilboa-Schechtman, E., & Feldman, R. (2017). Maternal depression and anxiety, social synchrony, and infant regulation of negative and positive emotions. *Emotion*, *17*(1), 11–27. https://doi.org/10.1037/emo0000204

Gunnery, S. D., & Ruben, M. A. (2016). Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cognition & Emotion*, *30*(3), 501–515. https://doi.org/10.1080/02699931.2015.1018817

Jeni, L. A., 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

Lucey, S., Ashraf, A. B., & Cohn, J. F. (2007). Investigating spontaneous facial action recognition through AAM representations of the face. In K. Delac & M. Grgic (Eds.), *Face recognition* (Vol. 2). IntechOpen. https://doi.org/10.5772/4841

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. https://doi.org/10.1371/journal.pone.0080161

Mattson, W. I., Ekas, N. V., Lambert, B., Tronick, E., Lester, B. M., & Messinger, D. S. (2013). Emotional expression and heart rate in high-risk infants during the face-to-face/still-face. *Infant Behavior and Development*, *36*(4), 776–785. https://doi.org/10.1016/j.infbeh.2012.11.009

Mesman, J., van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review*, *29*(2), 120–162. https://doi.org/10.1016/j.dr.2009.02.001

Messinger, D. S. (2002). Positive and negative: Infant facial expressions and emotions. *Current Directions in Psychological Science*, *11*(1), 1–6.

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. https://doi.org/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. https://doi.org/10.1037/a0026498

Mitsven, S. G., Messinger, D. S., Moffitt, J., & Ahn, Y. A. (2020). Infant emotional development. In J. Lockman & C. Tamis-Lemonda (Eds.), *Handbook of infant development* (pp. 748–782). Cambridge University Press.

Montirosso, R., Provenzi, L., Tavian, D., Missaglia, S., Raggi, M. E., & Borgatti, R. (2016). COMTval158met polymorphism is associated with behavioral response and physiologic reactivity to socio-emotional stress in 4-month-old infants. *Infant Behavior and Development*, *45*, 71–82. https://doi.org/10.1016/j.infbeh.2016.09.006

Montirosso, R., Provenzi, L., Tronick, E., Morandi, F., Reni, G., & Borgatti, R. (2014). Vagal tone as a biomarker of long-term memory for a stressful social event at 4 months. *Developmental Psychobiology*, *56*(7), 1564–1574. https://doi.org/10.1002/dev.21251

Montirosso, R., Tronick, E., Morandi, F., Ciceri, F., & Borgatti, R. (2013). Four-month-old infants' long-term memory for a stressful social event. *PLoS One*, *8*(12), e82277. https://doi.org/10.1371/journal.pone.0082277

Moore, G. A., Cohn, J. F., & Campbell, S. B. (2001). Infant affective responses to mother's still face at 6 months differentially predict externalizing and internalizing behaviors at 18 months. *Developmental Psychology*, *37*(5), 706–714. https://doi.org/10.1037//0012-1649.37.5.706

Onal Ertuğrul, I., Ahn, Y. A., Bilalpur, M., Messinger, D. S., Speltz, M. L., & Cohn, J. F. (2022). Infant AFAR: Automated facial action recognition in infants. *Behavior Research Methods*, *55*(3), 1024–1035. https://doi.org/10.3758/s13428-022-01863-y

Onal Ertuğrul, I., Cohn, J. F., Jeni, L. A., Zhang, Z., Yin, L., & Ji, Q. (2019). Cross-domain AU detection: Domains, learning approaches, and measures. In *Proceedings of the 2019 14th IEEE international conference on automatic face and gesture recognition (FG 2019)* (pp. 1–8). https://doi.org/10.1109/FG.2019.8756543

Onal Ertuğrul, I., Jeni, L. A., Ding, W., & Cohn, J. F. (2019). AFAR: A deep learning based tool for automated facial affect recognition. In *Proceedings of the international conference on automatic face and gesture recognition, IEEE international conference on automatic face and gesture recognition*. https://doi.org/10.1109/fg.2019.8756623

Oster, H. (2003). Emotion in the infant's face. *Annals of the New York Academy of Sciences*, *1000*(1), 197–204. https://doi.org/10.1196/annals.1280.024

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

Provenzi, L., Fumagalli, M., Bernasconi, F., Sirgiovanni, I., Morandi, F., Borgatti, R., & Montirosso, R. (2017). Very preterm and full-term infants' response to socio-emotional stress: The role of postnatal maternal bonding. *Infancy*, *22*(5), 695–712. https://doi.org/10.1111/infa.12175

Reck, C., Müller, M., Tietz, A., & Möhler, E. (2013). Infant distress to novelty is associated with maternal anxiety disorder and especially with maternal avoidance behavior. *Journal of Anxiety Disorders*, *27*(4), 404–412. https://doi.org/10.1016/j.janxdis.2013.03.009

Stockdale, L. A., Porter, C. L., Coyne, S. M., Essig, L. W., Booth, M., Keenan-Kroff, S., & Schvaneveldt, E. (2020). Infants' response to a mobile phone modified still-face paradigm: Links to maternal behaviors and beliefs regarding technoference. *Infancy*, *25*(5), 571–592. https://doi.org/10.1111/infa.12342

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Weinberg, M. K., Olson, K. L., Beeghly, M., & Tronick, E. Z. (2006). Making up is hard to do, especially for mothers with high levels of depressive symptoms and their infant sons. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*, 670–683, 16790002. https://doi.org/10.1111/j.1469-7610.2005.01545.x

Yaari, M., Rotzak, N. L., Mankuta, D., Harel-Gadassi, A., Friedlander, E., Eventov-Friedman, S., Bar-Oz, B., Zucker, D., Shinar, O., & Yirmiya, N. (2018). Preterm-infant emotion regulation during the still-face interaction. *Infant Behavior and Development*, *52*, 56–65. https://doi.org/10.1016/j.infbeh.2018.05.008

Yato, Y., Kawai, M., Negayama, K., Sogon, S., Tomiwa, K., & Yamamoto, H. (2008). Infant responses to maternal still-face at 4 and 9 months. *Infant Behavior and Development*, *31*, 570–577. https://doi.org/10.1016/j.infbeh.2008.07.008

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ahn, Y. A., Önal Ertuğrul, I., Chow, S.-M., Cohn, J. F., & Messinger, D. S. (2023). Automated measurement of infant and mother Duchenne facial expressions in the Face-to-Face/Still-Face. *Infancy*, 1–20. https://doi.org/10.1111/infa.12556