Infant Crying Predicts Real-Time Fluctuations in Maternal Mental Health in Ecologically Valid Home Settings

Kaya de Barbaro[1], Megan Micheletti[1], Xuewen Yao[2], Priyanka Khante[2], Mckensey Johnson[1], and Sherryl Goodman[3]

[1] Department of Psychology, The University of Texas at Austin
[2] Department of Electrical and Computer Engineering, The University of Texas at Austin
[3] Department of Psychology, Emory University

Exposure to infant crying is a well-established predictor of mothers' mental health. However, this association may reflect many potential mechanisms. Capturing dynamic fluctuations in mothers' states simultaneously with caregiving experiences is necessary to identify the real-time processes influencing mental health. In this study, we leveraged ecological momentary assessments (EMAs) and infant-worn audio recorders to capture variability in mothers' mental health symptoms and their exposure to infant crying over one week in a racially and socio-economically diverse urban North-American sample ($N = 53$). We use multilevel modeling to characterize within- and between-person effects of crying on maternal negative affect and symptoms of depression and anxiety. Within participants, when infants cried more than average in the 10 min, 1 hr, and 8 hr prior to an EMA report, mothers' negative affect subsequently increased, controlling for mean levels of infant crying. In contrast to findings from laboratory studies, in everyday settings crying exposure did not immediately increase feelings of depression. Only when crying was above average for 8 hr prior to EMA did mothers report increases in subsequent depression symptoms, suggesting that the effects of crying on maternal mental health take *hours* to unfold in ecologically valid home settings. Between participants, mothers of infants who cried more on average did not report higher negative affect or symptoms of depression or anxiety. Overall, our findings reveal that crying exposure dynamically influences maternal negative affect and depression but not anxiety in ecologically valid real-world settings.

Public Significance Statement
This study used cellphone surveys and infant-worn audio recorders to reveal that exposure to infant crying increases maternal negative affect and symptoms of depression, but not symptoms of anxiety, in real-world settings. In contrast to findings from laboratory experiments, we found that the effects of crying on maternal depression symptoms take *hours* (rather than minutes) to emerge. Studying mother–infant interactions in everyday home settings is necessary to understand and support maternal mental health.

Keywords: maternal mental health, infant crying, LENA wearable audio recorders, ecological momentary assessment, multilevel modeling

Supplemental materials: https://doi.org/10.1037/dev0001530.supp

Infant crying is a salient feature of a mother's postpartum environment and has well-established associations with maternal mental health (Young et al., 2012). Crying duration peaks with nearly 3 hr of crying per day when infants are four to six weeks of age and stabilizes to approximately 1 hr per day by three months (Vermillet et al., 2022; Wolke et al., 2017). However, there are

large individual differences in daily crying duration (Kurth et al., 2011; St James-Roberts & Plewis, 1996) and 5%–25% of infants are identified with colic, or excessive crying (Lucassen et al., 2001; Wolke et al., 2017). Mothers who report their infants cry excessively also report experiencing higher levels of caregiver frustration (Barr et al., 2014; Fujiwara et al., 2011) and emotional distress (Miller et al., 1993), as well as higher concurrent depression and anxiety symptoms (Petzoldt, 2018). Furthermore, mothers of infants with a temperament profile which includes high levels of irritability and crying also exhibit higher levels of depression and anxiety symptoms (Britton, 2011).

Theoretical models propose that crying exposure influences maternal mental health directly, at the level of individual processes (Kurth et al., 2011; Papoušek & von Hofacker, 1995). For example, Dix's Affective Organization of Parenting Model proposes that infant crying behaviors can activate immediate emotional reactions in caregivers (Dix, 1991). However, few studies examine how individual mothers' mental health may fluctuate over real time, and we found no published work that examined the real-world, real-time individual processes by which crying exposure influences maternal mental health. This is critical as there are many potential mechanisms that could lead to group-level associations, that is, associations *between* individuals versus *within* individuals. Researchers increasingly recognize the need to disaggregate within- and between-persons effects to examine processes of change central to psychology (Berry & Willoughby, 2017; Curran & Bauer, 2011; Hoffman & Stawski, 2009). Infant crying may be associated broadly with elevated maternal distress due to stable or "trait-like" differences that increase risks for both factors (Shackman et al., 2016). For example, maternal anxiety and depression may be associated with caregiving behaviors thought to intensify infant crying, such as maternal intrusiveness or inefficiency of soothing interactions (Papoušek & von Hofacker, 1995), respectively. Thus, although mothers whose infants cry more may have more symptoms of depression or anxiety, by virtue of examining differences between mothers, whether cross-sectionally or longitudinally, extant studies do not permit inferences about the processes leading to these associations.

Experimental crying exposures in a laboratory setting lend initial evidence to the immediate causal effects of crying on maternal affect. Women exposed to 10-min audio recordings of uncontrollable infant crying reported immediate increases in negative affect, including feelings of irritability, guilt, and hopelessness (Barr et al., 2014; Bruning & McMahon, 2009). However, mothers in Barr et al. (2014) were not exposed to their own infant's crying and the women in Bruning and McMahon (2009) were female undergraduates not mothers, both of which may not fully represent the association between maternal affect and infant crying. Additionally, these experimental paradigms distort crying in ways that are not representative of everyday crying exposures. For example, infant crying in everyday settings is more likely to occur in intense but relatively short bursts—lasting on average 3.6 s (Irwin, 2003)—compared with a continuous 10 min. Additionally, mothers' actions can and often do help to soothe infant crying (Hubbard & van IJzendoorn, 1991), meaning that "everyday" crying is not necessarily uncontrollable. Thus, exposure to ecologically valid crying may be less likely to immediately induce caregiver distress. It is also possible that lab-based exposures are too short to fully characterize the effect of crying exposure over the course of a day. Clinical observations suggest that persistent crying (i.e., crying that extends for many hours in a day) may be a particular risk factor for maternal distress (Papoušek & von Hofacker, 1995). Extended periods of crying accumulated over hours may have distinct effects on maternal affect relative to immediate or short-term effects. We found no published study that systematically characterized the impact of exposure to infant crying over minutes and hours on maternal mental health in an ecologically valid setting.

A daily diary or EMA studies are one approach that can be leveraged to access maternal mental health symptoms in the everyday context. In EMA studies, participants are repeatedly surveyed about their experiences on their mobile phones as they go about their lives. Relative to retrospective surveys in which participants are asked to report on their experiences from the past few days or weeks, EMAs can reduce recall bias, which is especially common in assessments of emotional experience (Trull & Ebner-Priemer, 2013). Critically, EMA studies also allow researchers to document the processes by which mood and emotion may fluctuate over time (Ebner-Priemer & Trull, 2009). Research designs that combine mood reports with objective measures of experience can be used to test theories of real-time effects of everyday contexts on mental health (Buhler-Wassmann & Hibel, 2021; Hamaker & Wichers, 2017; Shiffman et al., 2008). For example, time-varying measures of crying exposure can be time-locked to self-report surveys to assess lagged associations between infant crying and maternal mental health. Such high-density multimodal datasets can test precisely to what extent and at what timescales dynamic fluctuations in cry exposure explain subsequent *within-person* changes in maternal affect and mental health.

Advances in wearable sensors, paired with algorithms that can automatically detect markers of interest, provide a relatively low-burden method for capturing large volumes of objective measures of context, including infant crying (de Barbaro, 2019; Micheletti et al., 2022). Researchers have successfully used sensors to assess objective measures of sleep, physical activity, and even geographic location as real-time predictors of mental health (Mohr et al., 2017). We build on this work by collecting automatically detected infant crying as a predictor of women's mental health postpartum. Parent reports are known to overestimate crying by a factor of three (Cabana et al., 2021) and maternal psychological distress has been shown to have a small but significant biasing effect on reports of child crying behaviors (Atella et al., 2003). As such, objective markers of infants' crying can reduce shared measurement bias between assessments of maternal mental health and this common caregiving experience. Additionally, by virtue of providing precise, quantitative data about the amount of crying experienced by caregivers in the minutes and hours prior to EMA mood reports, wearable sensors could be used to examine potentially distinct effects of naturally occurring crying immediately versus at longer timescales on mothers' mental health.

We considered objective markers of cry duration as independent, continuous predictors of maternal mental health. Most past work examined associations between maternal mental health and exposure to excessive infant crying, defined as more than 2 or 3 hr of crying per day, for two to three days per week over the course of multiple months, as reported by the mother (Wessel et al., 1954). However, inconsistencies between parent-reported and objectively detected amounts of infant crying mean it is unclear how to validly translate this threshold to objective cry data. Additionally, if crying exposure has continuous effects on women's mental health, this would

suggest that women could more broadly benefit from interventions aimed to either reduce infant crying or reduce mothers' crying-related distress. As such, reviews of this literature (Kurth et al., 2011) suggest that new studies should examine the continuous effects of crying on maternal mental health.

The Current Study

To understand the extent of naturally occurring fluctuations in women's mental health and the extent to which they can be predicted by recent crying exposure, we combined EMAs of maternal mental health with automatically detected markers of infant crying collected from audio recorders worn by infants in their naturalistic environments for up to 72 hr over the course of one week. Drawing from clinical observations (Papoušek & von Hofacker, 1995), we tested whether crying exposure had distinct effects when occurring immediately prior to an EMA report (10 min or 1 hr) versus when it was accumulated over multiple hours of the day (8 hr).

Study Hypotheses

Our first hypothesis was that mothers' mental health, indexed as negative affect and symptoms of depression and anxiety, would fluctuate over the course of the week. Second, we hypothesized that exposure to more infant crying would lead to mothers reporting higher negative affect, depression symptoms, and anxiety symptoms, both between- and within-persons. Third, we hypothesized that mothers who were exposed to more-than-average infant crying in recent minutes as well as persistently over the course of hours would report higher negative affect, depression symptoms, and anxiety symptoms. The alternative hypothesis was that crying exposure poses only immediate, short-term effects on mothers, which do not persist after one or more hours.

Method

Participants

We recruited a diverse population of mothers and infants via convenience sampling from the Austin, Texas, metropolitan area. We advertised the study using Facebook, online university event calendars, and fliers posted in local community health centers. Exclusion criteria included infants older than 12 months, infants with congenital birth defects or a genetic condition affecting development (e.g., Down syndrome), multiple birth infants, and families with a primary residence located over 20 miles from the research center.

Mother–infant dyads ($n = 87$) were recruited in the context of a broader study leveraging wearable sensors to measure mother and infant behavior in the home. After an initial piloting phase with $n = 28$ mothers who tested the survey platform and provided feedback on user experience and feasibility, $n = 59$ mothers received mobile surveys asking about their mental health. Three mothers were excluded because they did not answer any surveys. An additional three mothers were excluded because they responded to only two surveys. The final sample included $n = 53$ mother–infant dyads. Table 1 shows participant characteristics. Infants were on average 3.9 months of age ($SD = 2.2$, range $= 0.9$–10.6 months) and 53% were female; 57% of dyads were non-Hispanic white.

Table 1
Participant Characteristics (n $= 53$)

Characteristic	M (SD), range or N (%)
Infant	
Age, months	3.9 (2.2), 0.9–10.6
Gestational age, weeks	39 (2), 31–41
Female sex	29 (52%)
Mother	
Age, years	31(6), 18–43
Female sex	53 (100%)
Education	
High school or less	9 (17%)
Some college or trade school	14 (26%)
College	14 (26%)
Graduate school	16 (31%)
Race/ethnicity	
White, Hispanic	6 (11%)
White, Non-Hispanic	30 (57%)
Black	4 (8%)
Hispanic	10 (19%)
Asian	1 (2%)
Multiracial	2 (4%)
Employment	
Full-time	21 (40%)
Part-time	10 (19%)
Not employed out of the home	22 (41%)
Family status	
Married	38 (72%)
Divorced	1 (2%)
Single parent	5 (9%)
Living with a partner without marriage	9 (17%)
Household income	
Under $25k	9 (17%)
$25k–49k	8 (15%)
$50k–74k	10 (19%)
$75k–99k	9 (17%)
Over $100k	17 (32%)
Primary household language	
English	44 (83%)
Spanish	9 (17%)
First-time mothers	19 (36%)

Note. Overall column includes mean (SD), range or frequency count (%)

Procedure

The University of Texas at Austin's Institutional Review Board approved the study design and data collection (Protocol: 2017-06-0026). Data were collected from September 2018 to February 2020.

Introductory Session

A research assistant conducted a 90-min study introduction in each participant's home. Research assistants obtained written informed consent, configured devices, and collected baseline information from mothers and infants while they wore the devices. The study was conducted in English or Spanish depending on the household's primary language.

Home-Recording Session

Following the introductory session, research assistants instructed mothers on how to record 72 hr of data over the course of the following week (Figure 1). Mothers were asked to record one 48-hr shift

Figure 1
Study Protocol
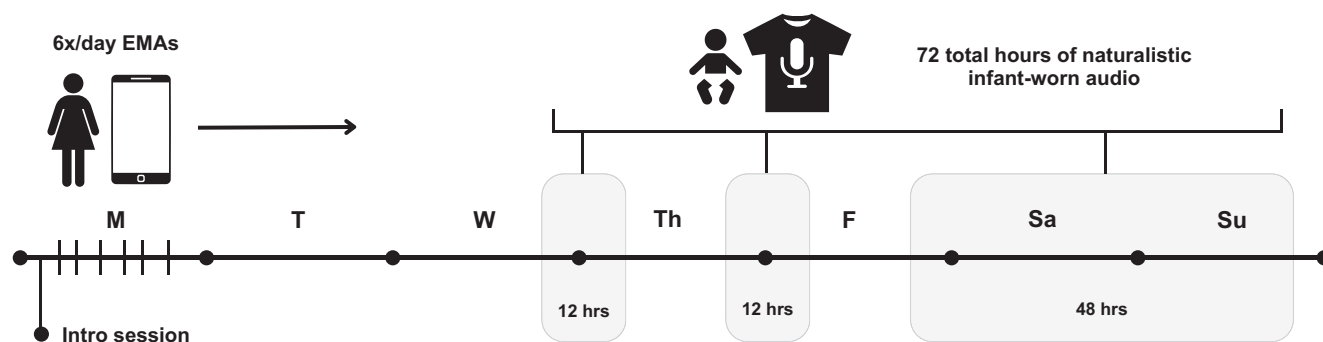
Note. Following a 90-min introductory session in the home, mothers completed six EMAs per day and infants wore audio recorders for a total of 72 hr over the course of seven days. EMA = ecological momentary assessment; audio = LENA recording device. Copyright by L. Gan for the baby icon, by Pixabay 420494 for the microphone icon, by Pixabay Clker-free-vector-images for the black t-shirt icon, by Pixabay MTZD for the mobile phone icon, and by Pixeden for the female icon, all via Canva.com. Reprinted with permission.

during which they planned to be with their child, typically the weekend, and two 12-hr shifts overnight. We selected these recording windows because they were largely feasible for mothers working full-time out of the home, while still allowing for ample continuous device wear time. To maximize adherence to the recording paradigm, mothers selected their recording windows during the introductory session and were reminded to record during the week by text messages from a research assistant. Following the session, we compensated mothers $100 for participation and infants received a small toy.

Measures

Given the evidence that negative affect shifts dynamically throughout the day (Trull & Ebner-Priemer, 2013), we used EMAs to capture mothers' immediate reports of negative affect six times per day. Mothers' anxiety and depression symptoms were assessed once per day in the evening. Anxiety and depression symptoms are not commonly assessed with high-density EMA designs (see reviews by aan het Rot et al., 2012; Ebner-Priemer & Trull, 2009), likely because they are theorized to reflect more stable or slowly changing clinical states. Indeed, a recent study that collected morning and evening depression surveys for six weeks found that morning and evening depression symptom levels were highly correlated (Targum et al., 2021). As such, we collected depression and anxiety symptoms by EMA one time per day in the evening, between 19:10 and 22:00. All surveys were delivered to the mothers via a text message containing a link to a 2- to 3-min online Qualtrics survey (Hofmann & Patel, 2015). Signals were considered semi-random because they were delivered at least 2 hr apart, but at otherwise random times between 05:00–07:49, 07:50–10:39, 10:40–13:29, 13:30–16:19, 16:20–19:09, and 19:10–22:00.

Negative Affect

Mothers were asked to report momentary feelings of negative affect six times per day by responding to a subset of items on the Positive and Negative Affect Schedule (PANAS; Peeters et al., 2003; Watson et al., 1988). Mothers read each of the following words and responded on a scale from 1 (*not at all*) to 7 (*very*) to indicate how strongly they felt that affect in the present moment: tense,

anxious, guilty, irritable, and easily distracted. Following Peeters et al., (2006) we included the item "easily distracted" as their research suggests it is important to include activation/arousal dimensions (in addition to traditional pleasantness/unpleasantness dimensions) in measures of momentary negative affect because they capture clinically relevant aspects of *daily* emotional experience. To minimize the response burden of our dense EMA sampling schedule, we excluded three of the eight items from the original scale (irritated, restless, and agitated) based on their overlap with the remaining items. To obtain a single negative affect score, we summed mothers' responses at each EMA administration with scores ranging from 5 to 35; higher scores indicate higher levels of momentary negative affect. Between-person ($\omega = 0.93$) and within-person ($\omega = 0.77$) reliabilities were high.

Depression

Mothers reported daily levels of depression symptoms once per day in the evening. Depression was measured with the well-validated Patient Health Questionnaire-2 (PHQ-2; Kroenke et al., 2009), a two-item measure used to assess depression symptom severity. Mothers were prompted to "think about today" and on a scale from 0 (*very slightly or not at all*) to 4 (*extremely*) rate the extent to which they had felt "little interest or pleasure in doing things" and "down, depressed, or hopeless." Total scores ranged from 0 to 8, with higher scores indicating higher symptomatology. The anhedonia criterion for depression was defined as "little interest or pleasure in doing things" and the depressed mood criterion was represented by "down, depressed, or hopeless." Between-person reliability ($\omega = 0.88$) was higher than within-person reliability ($\omega = 0.32$), meaning that responses on the PHQ-2 were not consistent within individual measurement occasions. However, this is expected given that the two items capture different criteria for depression symptomatology.

Anxiety

Anxiety symptoms were measured once per day in the evening using the well-validated Generalized Anxiety Disorder 2-item (GAD-2; Kroenke et al., 2009). On a scale from 0 (*very slightly or*

not all) to 4 (*extremely*), mothers rated the extent to which they were "feeling nervous, anxious, or on edge" and "not being able to stop or control worrying" thinking about today. Scores ranged from 0 to 8 where higher scores reflect more anxiety symptoms. Between-person ($\omega = 0.89$) and within-person ($\omega = 0.68$) reliabilities were high.

Infant Crying

Objective measures of infant crying duration were obtained using Language EnviroNment Analysis (LENA) technology (Xu et al., 2009). The LENA device is a small, child-safe recorder worn in a vest over the infant's clothing. The device records the infant's voice as well as other voices and sounds in the environment up to 10 feet away. We collaborated with LENA technology to increase the device's storage capacity so it can record for a continuous 24 hr. During infant bathing and sleep, mothers removed the vest but kept it nearby and recording to continue capturing ambulatory audio.

To capture infant crying in recent minutes as well as persistent crying over the course of hours, we calculated time-varying measures of crying across three different time intervals preceding each survey response for each mother: 10 min, 1 hr, and 8 hr. We selected these intervals a priori because they allowed us to examine the short-term effects of crying (similar to lab-based experiments where 10-min audio recordings were played for mothers) as well as explore a potential long-term effect of crying over the course of the day, where persistent crying may be more distressing. We chose 8 hr to maximize usable data within our recording paradigm; we could not evaluate the effect of crying in the past 24 hr on negative affect and symptoms of depression and anxiety because the LENA device was often not worn continuously for a full day prior to each survey response. We also considered including crying in the 3 hr preceding survey response as an additional predictor, but it was highly correlated to 1 hr ($r = .73$) and 8 hr ($r = .80$), so we did not include it given our relatively small sample size.

Infants had to be wearing LENA for at least 70% of the intended time interval (i.e., in the previous 10 min, 1 hr, or 8 hr) for their data to be included in each analysis. Crying values were normalized per minute and scaled up to the full interval. For example, if an infant wore LENA for 7 of the last 8 hr, cry duration was normalized per minute and scaled up to the full 8 hr.

Any nonverbal vocal expression of negative emotion, including cries, fusses, whines, and whimpers, were automatically identified by LENA as "crying" (Fields-Olivieri & Cole, 2019). As LENA's cry algorithm has been shown to underestimate crying time relative to human annotations (Micheletti et al., 2022), we compared LENA's cry output to human annotations of cry for 40% of our participants. Trained human coders annotated 24-hour recordings for $n = 22$ participants quasi-randomly selected to represent the age distribution of our overall sample. A cry had to have a minimum duration of three seconds and fusses/whines/whimpers had no minimum duration (Gilkerson et al., 2017); all were combined and referred to as cries. Human coders achieved a Cohen's κ of 0.77 for detecting cry and non-cry episodes, which represents a moderate to strong level of agreement (McHugh, 2012). We found that LENA underestimated cry duration by a factor of three, missing up to 48 min of crying every 24 hr. Given this sizeable discrepancy, we estimated a linear regression model to predict human-annotated cry durations from LENA

outputs at each time interval. Then, we used the parameters of this regression to predict human-annotated crying from LENA output at each time interval (see Table 2 for regression equations). In effect, we multiplied all original durations by approximately a factor of three which resulted in cry estimates that more closely approximate the gold-standard human-annotated values for each time frame. Converging evidence suggests that a correction factor should be applied to LENA crying data to compensate for its underestimation of child vocalizations (Cristia et al., 2021; Micheletti et al., 2022). This transformation is equivalent to changing from one unit to another and thus has no effect on our results. The main benefit is that it provides a more accurate estimate of infant crying duration—that is, the transformed values are more representative of the true durations of crying as coded by trained research assistants.

In addition to crying duration, we also considered evaluating cry frequency, or the number of cries in each time interval. However, cry frequency was highly correlated with cry duration at 10 min ($r = 0.92$), 1 hr ($r = 0.90$), and 8 hr ($r = 0.72$). To minimize our total number of comparisons, and because much of the existing literature focuses on total crying time, we evaluated cry duration only (Kurth et al., 2011).

Data Analysis

All data analyses were conducted in R (4.2.1) and R Studio (v2022.07.1). First, we calculated single rater, one-way random effects intra-class correlation coefficients (Bliese, 2000) to describe the percentage of total variance in negative affect, depression, and anxiety symptom scores accounted for by between- and within-person differences. We also calculated correlations between negative affect, depression symptoms, and anxiety symptoms to evaluate the extent to which our outcomes were associated in real-time.

Given the skewness of our observed data, we calculated which transformation (log, square root, or exponential) would result in the most normally distributed data for each predictor and outcome variable (Peterson & Cavanaugh, 2020). We then applied the best normalizing transformation to each predictor and outcome variable to prepare for multilevel regression analyses (see Table 3). Next, we used two-level multilevel modeling with maximum likelihood estimation and robust standard errors to account for repeated observations nested within persons. Random intercepts were allowed for Level 2 (participant) membership. We tested infant crying in the prior 10 min, 1 hr, and 8 hr as a predictor of subsequent negative affect, depression symptoms, and anxiety symptoms. Each model included within-person variability in crying and person-centered mean crying calculated across all observations.

Table 2

Correlations and Linear Regressions Between LENA and Human-Annotated Crying by Time Interval

Time interval preceding EMA response	R	Regression equation
10 min	0.68	$y = 2.8268x + 0.2253$
1 hr	0.72	$y = 2.9304x + 0.6545$
8 hr	0.73	$y = 2.8488x + 10.743$

Note. Linear regression where $y =$ human-annotated crying and $x =$ LENA-detected crying. EMA = ecological momentary assessment; LENA = Language EnviroNment Analysis.

Table 3

Estimated Normality Statistic by Transformation Type

Variable	log	Square root	Exponential
Negative affect	5.5526	**5.2018**	178.2202
Depression symptoms	15.2903	**15.1315**	25.9884
Anxiety symptoms	7.2528	**7.1515**	24.6302
10 min cry duration	**30.0647**	32.3507	43.1608
1 hr cry duration	**11.0131**	11.341	13.88
8 hr cry duration	11.247	**6.4279**	56.496

Note. Estimated normality statistic is measured by Pearson P/degrees of freedom, where lower values reflect a more normal distribution. The best-normalizing transformation (in bold) was applied to each variable prior to multilevel modeling.

Infant age (in weeks), first-time mothers (1 = *yes*; 0 = *no*), and maternal education (high school or less, some college or trade school, college, graduate school) were selected a priori to be included as covariates in all models. We included infant age because crying in a younger versus older infant could have different effects on maternal mood and mental health, for example, as older infants spend less time crying per day (Vermillet et al., 2022; Wolke et al., 2017), crying more than usual for an older infant may be more unexpected and thus more distressing. We also controlled for first-time mothers, considering that first-time mothers may have less caregiving experience and may be more affected by their infant's crying. For example, first-time mothers have been shown to be less confident in interpreting their infant's cues and more likely to report crying as a problem relative to multiparous mothers (Canivet et al., 2005; Drummond et al., 1993). Lastly, we included maternal education to serve as a marker of socioeconomic status, particularly because lower maternal education has been associated with higher parenting stress (Fang et al., 2022), and higher parenting stress may contribute to a reduced reactivity to an infant's cries (Kim et al., 2020).

Given our relatively small sample size and the issue of multicollinearity we did not include all remaining demographic variables in our models. Thus, for the remaining demographic variables we took a bottom-up approach and tested the association between all potential covariates via regression or chi-squared test. We found that maternal education was significantly related to income (χ^2 = 1,286.9, $p < .001$), family status (χ^2 = 559.01, $p < .001$), employment (χ^2 = 322.58, $p < .001$), and maternal age (B = 2.5, $p < .001$). Specifically, higher maternal education was associated with higher family income, more married couples, more mothers working full-time, and older mothers, so we included only maternal education as a marker of socioeconomic status in each model.

A power analysis based on 1,000 Monte Carlo simulations showed that this study was powered at an average of 87% to detect medium-sized effects of 0.5 and powered at an average of 52% to detect small-sized effects of 0.2 with an alpha level of .05 (Green & MacLeod, 2016). Given the normalizing transformations applied to our data, estimated beta values cannot be directly interpreted. Thus, to facilitate comparison across predictors, for all models we report standardized coefficients (β), equivalent to the change in the outcome variable associated with a standard deviation change in the predictor variable (Snijders & Berkhof, 2008).

This study was not preregistered. Data and study materials from this study are not publicly available. Syntax for our linear mixed effects models is presented in the online supplemental materials.

Results

Mothers responded to an average of 33 of 42 planned EMA survey signals (SD = 11, median = 33, range = 10–52). Considering only the last EMA of the day (where depression and anxiety symptoms were reported), mothers responded to an average of six of seven planned survey signals (SD = 2, range = 0–10). Out of the scheduled 72 hr of device wear time, infants wore the LENA for a mean 71 hr (SD = 17, range = 17–103). As common in ambulatory assessment research, we observed large ranges in the amount of EMA and LENA data collected per participant. However, EMA-audio pairs were only included in each analysis if there was sufficient overlap in the timing between the EMA and the cry detection windows (i.e., at least 70% of audio available for each time-locked window). Note that due to technical difficulties, we scheduled an additional one to three days of data collection to obtain concurrent LENA and EMA data for n = 11 participants. This means that some participants completed surveys for more than seven days, which accounts for the high end of some ranges.

Table 4 shows descriptive summaries of EMA responses and infant crying at the person-centered mean level. Given that the PHQ-2 and GAD-2 were validated on a retrospective report of the past two weeks, while we ask about the current day, we cannot compare clinical cutoffs directly. However, 35 mothers (66%) reported a PHQ-2 score greater than zero on at least one occasion, endorsing symptoms "a little," while 18 mothers (34%) reported a three or higher on at least one occasion, indicating symptoms "quite a bit" to "extremely." For the GAD-2, 44 mothers (83%) reported a score greater than zero at least once and 28 mothers (53%) reported a three or higher at least once.

After correcting for skewness, we found that reports of negative affect were correlated with depression symptoms at $r = 0.47$ and anxiety symptoms at $r = 0.65$. Symptoms of depression and anxiety were also correlated at $r = 0.56$. This suggests that although there is some shared variance in our measures, as is expected given their high rates of co-occurrence (Brown et al., 2001), our main outcome measures do not entirely overlap and therefore additional variance in scores could be explained by exposure to infant crying.

Intraclass Correlation Coefficients

Variance attributable to between-person differences was 0.62 for negative affect, 0.63 for depression symptoms, and 0.53 for anxiety symptoms. The inverse of these values is attributable to within-person differences. Figure 2 shows the range and variability of

Table 4

Descriptive Statistics on Survey Responses and Infant Crying Durations at the Person-Centered Mean Level

Variable	n	M (SD)	Range
EMA survey			
Negative affect	53	11.5 (4.6)	5.6–24.4
Depression symptoms	52	0.9 (1.3)	0–5.3
Anxiety symptoms	52	1.4 (1.4)	0–6
Infant crying (min)			
10 min	45	0.4 (0.2)	0.2–1.0
1 hr	45	2.1 (0.7)	1.3–4.9
8 hr	45	13.0 (4.0)	6.2–23.4

Figure 2
Within-Person Variability in Maternal Negative Affect, Depression and Anxiety Symptoms, and Infant Crying
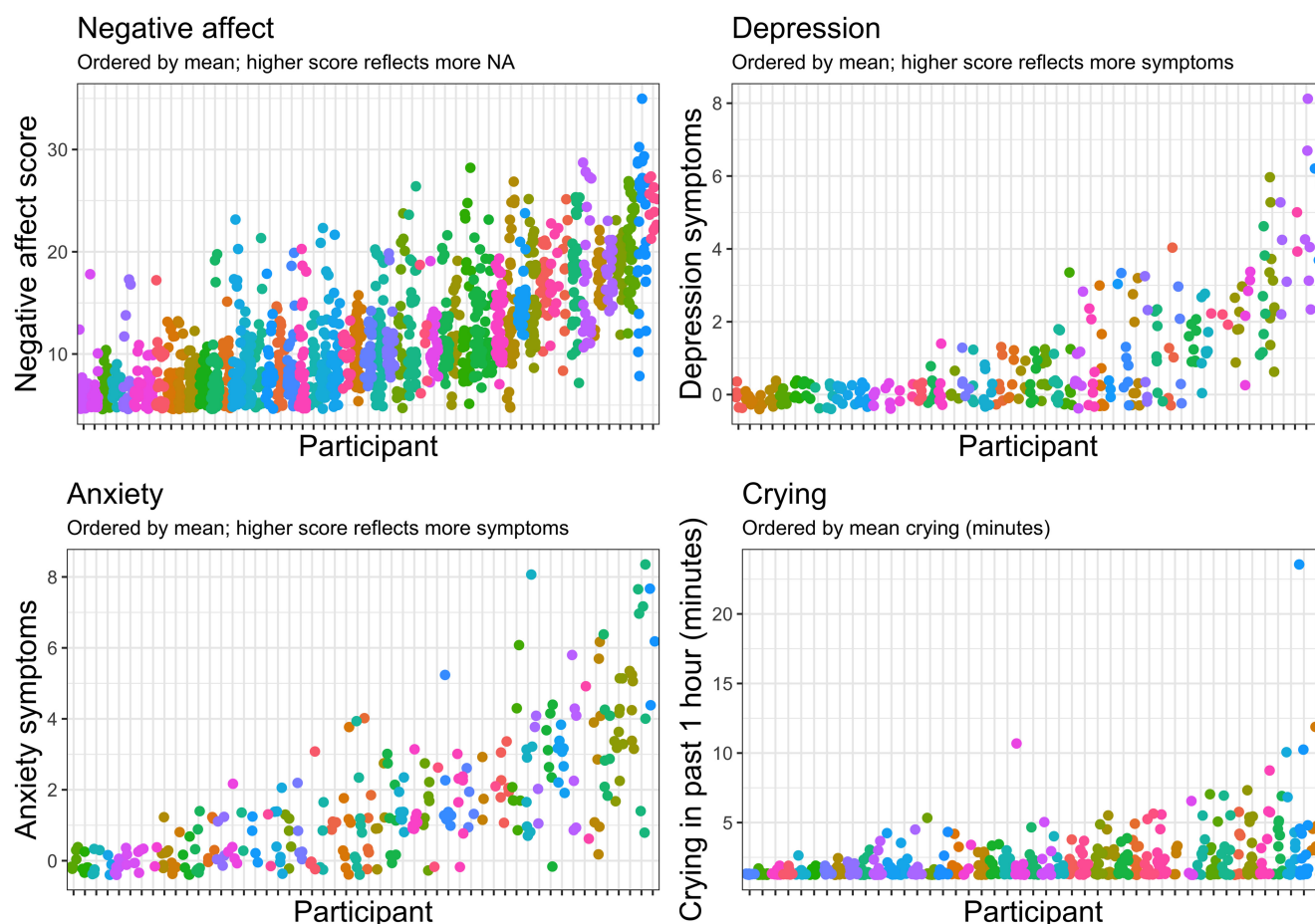
Note. Within-person variability in negative affect, 38%; depression symptoms, 37%; anxiety symptoms, 47%; and crying duration in the past 1 hr, 93%. Each column represents one participant's data, reflecting either raw mental health scores from each completed survey or crying detected in the 1 hr prior to each completed survey. Scores are jittered and participants are ordered from left to right by mean level to illustrate within-person variability. See the online article for the color version of this figure.

EMA responses and crying durations at 1 hr between and within persons. For example, there is more within-person variability in anxiety symptoms (0.47) than depression symptoms (0.37), meaning that mothers in our sample showed more day-to-day variability in their reports of anxiety symptoms compared to depression symptoms.

Multi-Level Modeling

Each model included the within-person variability in crying and the person-centered mean crying calculated across all observations. Separating crying exposure at each occasion into these two variables allows us to distinguish whether mothers' mental health symptoms are associated with *between-persons* differences in mean crying exposure (e.g., mothers of infants who on average tend to cry more vs. less) or whether mothers' mental health symptoms shift as a function of *within-person* changes in crying exposure (i.e., when a mother is exposed to more or less crying than usual, for that mother). Within-person deviations are equivalent

to centering each participant's crying exposure at zero to examine how positive versus negative deviations from their own mean continuously predict our outcomes. Table 5 summarizes model results.

Negative Affect

Within-person, mothers reported significantly higher negative affect on the occasions when their infant cried more-than-average in the 10 min ($b = .102$, $p < .001$), 1 hr ($b = .138$, $p < .001$), and 8 hr ($b = .074$, $p = .006$) prior to completion of individual mood reports. That is, on occasions when mothers were exposed to more infant crying than usual in the prior 10 min, 1 hr, and 8 hr, they reported significantly higher negative affect as compared with occasions when their infants cried less. Ratings of negative affect were not related to infants' mean crying duration in the 10 min, 1 hr, or 8 hr prior to EMA response. None of our covariates significantly predicted negative affect across models.

Table 5

Standardized Beta Coefficients From Linear Mixed Models Predicting Momentary Reports of Mental Health From Infant Crying at Varying Timescales

Infant crying	Negative affect				Depression symptoms				Anxiety symptoms			
	n	Beta	CI	*p*	*n*	Beta	CI	*p*	*n*	Beta	CI	*p*
Within-person variation												
10 min	44	0.10***	[0.18, 0.55]	<.001	43	0.11	[−0.20, 0.62]	.323	43	−0.01	[−0.58, 0.52]	.881
1 hr	44	0.14***	[0.35, 0.78]	<.001	42	0.11	[−0.18, 0.70]	.248	42	0.02	[−0.38, 0.50]	.800
8 hr	44	0.07**	[0.02, 0.10]	.006	36	0.35*	[0.02, 0.29]	.027	36	0.18	[−0.05, 0.23]	.236
Person-centered mean												
10 min	44	−0.02	[−2.55, 2.20]	.884	43	0.01	[−1.33, 1.39]	.964	43	−0.01	[−1.97, 1.87]	.955
1 hr	44	0.17	[−0.81, 3.74]	.216	42	0.12	[−0.71, 1.86]	.385	42	0.23	[−0.46, 3.14]	.154
8 hr	44	0.02	[−0.30, 0.36]	.859	36	0.06	[−0.17, 0.25]	.708	36	0.09	[−0.30, 0.46]	.585

Note. Within each model, we included the within-person variation and person-centered mean as predictors. Infant age, first-time mothers, and maternal education were included as covariates in all models.
*$p < .05$.   **$p < .01$.   ***$p < .001$.

Depression Symptoms

Within-person, on occasions that infants cried more than average in the 8 hr prior to the evening EMA, mothers reported higher ratings of depression symptoms relative to occasions when their infants cried less over the prior 8 hr ($b = .354$, $p = .027$). In this model, we also found that when exposed to the same amount of crying, mothers with some college reported higher depression symptoms than mothers with a college degree when exposed to more than average crying in the past 8 hr ($p = .015$). Within-person changes in crying in the 10 min or 1 hr prior to the evening EMA did not predict depression symptoms. Between-person, infants' mean crying, relative to other infants, did not predict symptoms of maternal depression at 10 min, 1 hr, or 8 hr.

As the PHQ-2 is comprised of two items that capture different criteria for depression—anhedonia ("little interest or pleasure in doing things") and low mood ("sad, depressed, or hopeless")—we conducted a post hoc analysis to evaluate if one item was driving our within-person results at the 8-hr timescale. We found that within-person increases in infant crying in the 8 hr prior to the evening EMA predicted higher reports of anhedonia relative to occasions in which infants cried less ($b = .236$, $p = .021$), but not low mood ($b = .049$, $p = .558$).

Anxiety Symptoms

No within-person changes in crying predicted daily symptoms of anxiety. Similarly, symptoms of anxiety were not predicted by infants who cried on average more in the 10 min, 1 hr, or 8 hr prior to EMA response.

Discussion

To observe the real-time processes by which crying exposure influences maternal mental health in the everyday context, we used objective markers of infant crying to predict in-the-moment reports of mothers' negative affect and symptoms of depression and anxiety over the course of one week. Overall, our results indicated that recent within-person increases in exposure to infant crying predicted higher maternal negative affect and depression symptoms, controlling for mean levels of our predictors. We also observed that mean levels of crying exposure did not predict mothers' reports of negative affect or symptoms of depression or anxiety. Our findings

reveal a dynamic within-person relationship between crying exposure and maternal mental health, providing the first evidence that infant crying influences mother's mental health at the level of the individual process in an ecologically valid setting.

Ecological momentary assessments indicated that mothers' negative affect and symptoms of depression and anxiety were not static; they fluctuated over the course of the week with 38% to 47% of variance attributable to within-person change. This meaningful within-person variability in mental health symptoms is broadly consistent with findings from the EMA literature (Ebner-Priemer & Trull, 2009), however, this is one of the first studies to examine such variability in a community sample of postpartum women (see also Bowen et al., 2012; Mendez et al., 2019).

We also found that dynamic fluctuations in infant crying, controlling for mean levels of crying, predicted mothers' feelings of negative affect and depression symptoms with distinct patterns across timescales. Variability in crying predicted negative affect at all time intervals, indicating that when mothers were exposed to more-than-average crying they subsequently felt worse mood immediately (at 10 min and 1 hr) as well as throughout the day (at 8 hr), relative to occasions when their infants cried less. Thus, we extend findings from the laboratory (Barr et al., 2014; Bruning & McMahon, 2009) to an everyday caregiving setting, that crying exposure can lead to increases in negative affect.

Additionally, our findings indicate distinct effects of crying over longer relative to shorter timescales. Within-person exposure to crying was associated with increases in depression symptoms, but only when crying was higher-than-average summing across the prior *8 hr* (i.e., when crying was more persistent than usual for the entire day). This effect was significant only for the anhedonia item of the PHQ-2, not low mood, which suggests that exposure to higher-than-average crying over the course of the day may manifest in postpartum women as more anhedonia (e.g., losing interest or pleasure in activities) versus a depressed mood state (e.g., feeling sad, hopeless, empty, or overwhelmed). We also found that within person, higher amounts of crying for shorter periods of time, that is for 10 min or 1 hr—even when above average—did not appear to increase mothers' feelings of depression relative to when their infants cried less. It is possible that mothers can cope or self-regulate when crying is above average for short periods of time, but as this time increases, mothers' resources become depleted and we begin to observe effects on mental health, similar

to the cumulative negative effects of stress on mental health over prolonged periods of time (Mcewen, 2004). These findings are not possible to obtain with traditional laboratory studies in which it is practically infeasible to manipulate exposure to crying for hours at a time.

More broadly, reports of depression symptom severity were less labile relative to negative affect, perhaps reflecting underlying differences in the process and rate at which affect and mood vary within individuals as a result of recent caregiving experiences. Of note, the null effects on depression and anxiety at shorter timescales may be the result of these reports being collected only once per day (vs. negative affect six times per day). As such, there may not have been enough samples at the 10 min and 1 hr timescales to detect results, a point we explore further in the limitations section.

Our between-person findings revealed that mean levels of crying exposure in the previous 10 min, 1 hr, and 8 hr did not predict mothers' reports of negative affect or symptoms of depression or anxiety. This is in contrast to previous findings that show longer and more frequent infant crying is associated with postpartum depression and anxiety (Britton, 2011; Petzoldt et al., 2016). However, these studies often relied on parent reports of infant crying, which may have led to shared measurement bias, and used thresholds for excessive or persistent infant crying, which our nonclinical, community sample of infants may not have met. Of note, a large systematic review of longitudinal studies of maternal depression, anxiety, and infant crying (Petzoldt, 2018) reported that excessive crying predicts future depression but not anxiety, which is consistent with our findings that within-person exposure to more crying is associated with real-time and same-day increases in negative affect and depression symptoms, but not anxiety symptoms, at the level of individual processes.

We also note that two models revealed a significant effect of maternal education, such that women with some college education exhibited more symptoms of depression, and specifically reports of anhedonia, compared with women with a college degree when exposed to more than average crying in the past 8 hr. This speaks to a possible risk factor associated with lower levels of maternal education, which is consistent with existing literature (Britton, 2008; Wesselhoeft et al., 2020). However, we interpret these results with caution as we find differences among women who all received some college education or higher; further research with a sample more diverse in levels of education is warranted.

Strengths and Limitations

The main strengths of this study are its use of a high-density, multimodal dataset to capture crying exposure and measures of affect and mood in real time, thus minimizing recall or reporting bias, as well as utilizing multilevel modeling to disaggregate and test hypotheses about between- and within-person effects. The simultaneous evaluation of negative affect, depression, and anxiety also facilitated our ability to examine the specificity of mother's mental health experiences with respect to crying exposure at different timescales. We also controlled for infant age, being a first-time mother, and maternal education in our statistical models. As such, our results are unlikely to simply reflect a mother's level of experience with a crying infant or her level of education, or only extend to a specific developmental window in infancy. Finally, the community-based nature of our sample supports the generalization of our findings to the broader population.

The present study also had several limitations. Participants were obtained through convenience sampling, meaning the sample is not representative of any specific postpartum community, despite representing socio-economic and racial diversity (Table 1). Additionally, by relying on a community sample with minimal eligibility criteria, our findings may not generalize to women who are currently or at elevated risk for experiencing postpartum mood or anxiety disorders. As such, further research is required to extend these findings to women with current or elevated risks for the development of postpartum depression and anxiety, including history of stressful life experiences, traumatic birth, or prior depressive episodes (O'Hara & McCabe, 2013). Our two-item measures for depression and anxiety symptoms were also brief, which allowed us to ask them repeatedly via EMA over the course of one week, but they did not capture the full spectrum of depression and anxiety symptomatology. Further research is needed to extend these findings to additional aspects of mothers' mental health as well as confirm that our findings reflect different mechanisms unique to depression and anxiety versus a general psychopathology factor.

Further research is also needed to determine whether our results would generalize to a sample of infants reported to be excessive criers. We focused on cry duration over other features of crying that could differentially impact mental health, like crying pitch or amplitude (LaGasse et al., 2005) as well as the inconsolability of infant crying (Barr et al., 2014). It is possible that within- and between-person changes in other features of crying could reveal differential effects on mental health, for example, perhaps it is the inability to soothe an infant that relates more to maternal anxiety compared to crying duration.

Our study was also underpowered to detect small-sized effects (0.20), particularly for depression and anxiety measured once per day relative to negative affect measured six times per day. A series of 1,000 Monte Carlo simulations revealed we were powered to detect a small-sized effect of within-person variability in crying on negative affect at 62% for 10 min, 53% for 1 hr, and 100% for 8 hr. We were powered to detect a small-sized effect of within-person variability in crying on symptoms of depression at 26% for 10 min, 21% for 1 hr, and 82% for 8 hr; for anxiety symptoms, power calculations were 26% for 10 min, 18% for 1 hr, and 77% for 8 hr. This means that there could still be a true effect of crying exposure on feelings of depression and anxiety—particularly at the 10 min and 1 hr timescales—that this study could not detect, so our null results should be interpreted with caution.

Finally, our study focused on linear, group-wide, unidirectional analyses. However, bidirectional, person-specific, and nonlinear effects also likely play a role in the maintenance and exacerbation of mothers' mental health and well-being (Papoušek & von Hofacker, 1995; Sadeh et al., 2010). Future studies should use person-specific and cross-lagged effects models to characterize more individualized, bidirectional dynamics between daily experience and maternal mental health.

Implications and Future Work

This study suggests that infant crying is variable between and within infants, and this within-person variability in crying produces real-time effects on mothers' emotions in a community sample that was not selected on the basis of clinical levels of depression or anxiety. While the beneficial effects of social support on mental health

in postpartum women are well-documented (e.g., Beck, 2001), our study suggests that mothers in general—that is, the broad community of mothers, irrespective of their mental health history—could benefit from support in caring for their crying infant and/or developing helpful skills to cope with the demands of caregiving. Having someone to provide emotional support or instrumental support, like actively responding to a distressed infant, could reduce the risk of worse affect following cry exposure. In extreme cases, excessive infant crying can lead to shaking and abusive head trauma, or Shaken Baby Syndrome (SBS), which is most likely to occur between four and six weeks as infant crying peaks (Barr et al., 2006). Teaching mothers how to regulate their emotions and modify their expectations and beliefs around crying, for example, viewing crying as a normal part of infant behavior, has been shown to prevent SBS (Stewart et al., 2011), and may provide mothers with in-the-moment tools to cope with stressful demands. Our findings of within-person effects of crying on maternal mental health suggest that such coping tools may be universally helpful rather than only for high-risk families. Future work may investigate possible moderators of the effect of crying on mental health, like perceived social support, parenting self-efficacy, and parenting stress, as well as the specific strategies by which mothers effectively regulate their emotions in the presence of infant crying.

Conclusion

Overall, exposure to infant crying is not a static feature of mothers' experience but, rather, variable day-by-day and even minute-by-minute. We found that within-person increases in infant crying exposure (across minutes as well as hours) increased mothers' reported negative affect and symptoms of depression. Insofar as ambulatory sensing methods can provide more systematic access to time-varying markers of experience, they open new possibilities to reveal the real-time processes linking the structure of an individual's ecological context with their mental health. Additionally, our within-person findings provide evidence that interventions to modify risk factors such as infant crying and reactions to infant cries are likely to provide direct benefits to individual mothers.

References

aan het Rot, M., Hogenelst, K., & Schoevers, R. (2012). Mood disorders in everyday life: A systematic review of experience sampling and ecological momentary assessment studies. *Clinical Psychology Review*, 32(6), 510–523. https://doi.org/10.1016/j.cpr.2012.05.007

Atella, L. D., DiPietro, J. A., Smith, B. A., & St James-Roberts, I. (2003). More than meets the eye: Parental and infant contributors to maternal and paternal reports of early infant difficultness. *Parenting*, 3(4), 265–284. https://doi.org/10.1207/s15327922par0304_1

Barr, R. G., Fairbrother, N., Pauwels, J., Green, J., Chen, M., & Brant, R. (2014). Maternal frustration, emotional and behavioural responses to prolonged infant crying. *Infant Behavior and Development*, 37(4), 652–664. https://doi.org/10.1016/j.infbeh.2014.08.012

Barr, R. G., Trent, R. B., & Cross, J. (2006). Age-related incidence curve of hospitalized Shaken Baby Syndrome cases: Convergent evidence for crying as a trigger to shaking. *Child Abuse & Neglect*, 30(1), 7–16. https://doi.org/10.1016/j.chiabu.2005.06.009

Beck, C. T. (2001). Predictors of postpartum depression: An update. *Nursing Research*, 50(5), 275–285. https://doi.org/10.1097/00006199-200109000-00004

Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development*, 88(4), 1186–1206. https://doi.org/10.1111/cdev.12660

Bliese, P. (2000). Within-group agreement, non-independence, and reliability. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations. Foundations, extensions, and new directions* (pp. 249–381). San Francisco Jossey-Bass. https://www.kellogg.northwestern.edu/rc/workshops/mlm/Bliese_2000.pdf

Bowen, A., Bowen, R., Balbuena, L., & Muhajarine, N. (2012). Are pregnant and postpartum women moodier? Understanding perinatal mood instability. *Journal of Obstetrics and Gynaecology Canada*, 34(11), 1038–1042. https://doi.org/10.1016/S1701-2163(16)35433-0

Britton, J. R. (2008). Maternal anxiety: Course and antecedents during the early postpartum period. *Depression and Anxiety*, 25(9), 793–800. https://doi.org/10.1002/da.20325

Britton, J. R. (2011). Infant temperament and maternal anxiety and depressed mood in the early postpartum period. *Women & Health*, 51(1), 55–71. https://doi.org/10.1080/03630242.2011.540741

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, 110(4), 585–599. https://doi.org/10.1037//0021-843X.110.4.585

Bruning, S., & McMahon, C. (2009). The impact of infant crying on young women: A randomized controlled study. *Journal of Reproductive and Infant Psychology*, 27(2), 206–220. https://doi.org/10.1080/02646830802350856

Buhler-Wassmann, A. C., & Hibel, L. C. (2021). Studying caregiver-infant co-regulation in dynamic, diverse cultural contexts: A call to action. *Infant Behavior and Development*, 64, Article 101586. https://doi.org/10.1016/j.infbeh.2021.101586

Cabana, M. D., Wright, P., Scozzafava, I., Olarte, A., DiSabella, M., Liu, X., & Gelfand, A. A. (2021). Newborn daily crying time duration. *Journal of Pediatric Nursing*, 56, 35–37. https://doi.org/10.1016/j.pedn.2020.10.003

Canivet, C. A., Östergren, P.-O., Rosén, A.-S., Jakobsson, I. L., & Hagander, B. M. (2005). Infantile colic and the role of trait anxiety during pregnancy in relation to psychosocial and socioeconomic factors. *Scandinavian Journal of Public Health*, 33(1), 26–34. https://doi.org/10.1080/14034940410028316

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., Bunce, J., & Bergelson, E. (2021). A thorough evaluation of the Language Environment Analysis (LENA) system. *Behavior Research Methods*, 53(2), 467–486. https://doi.org/10.3758/s13428-020-01393-5

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental Psychobiology*, 61(3), 444–464. https://doi.org/10.1002/dev.21831

Dix, T. (1991). The affective organization of parenting: Adaptive and maladaptive processes. *Psychological Bulletin*, 110(1), 3–25. https://doi.org/10.1037/0033-2909.110.1.3

Drummond, J. E., McBride, M. L., & Wiebe, C. F. (1993). The development of mothers' understanding of infant crying. *Clinical Nursing Research*, 2(4), 396–410. https://doi.org/10.1177/105477389300200403

Ebner-Priemer, U. W., & Trull, T. J. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*, 21(4), 463–475. https://doi.org/10.1037/a0017075

Fang, Y., Luo, J., Boele, M., Windhorst, D., van Grieken, A., & Raat, H. (2022). Parent, child, and situational factors associated with parenting stress: A systematic review. *European Child & Adolescent Psychiatry*. Advance online publication. https://doi.org/10.1007/s00787-022-02027-1