

# Mothers' Tone of Voice Depends on the Nature of Infants' Transgressions

Audun Dahl, Briana R. Sherlock, Joseph J. Campos, and Frédéric E. Theunissen  
University of California, Berkeley

Emotional vocal signals are important ways of communicating norms to young infants. The second year is a period of increase in various forms of child transgressions, but also a period when infants have limited linguistic abilities. Two studies investigated the hypothesis that mothers respond with different vocal emotional tones to 3 types of child transgressions: moral (harming others), prudential (harming oneself), and pragmatic (creating inconvenience, e.g., by spilling) transgressions. We used a combination of naturalistic observation (Study 1) and experimental manipulation (Study 2) to record, code, and analyze maternal vocal responses to child transgressions. Both studies showed that mothers were more likely to use intense, angry vocalizations in response to moral transgressions, fearful vocalizations in response to prudential transgressions, comforting vocalizations in response to pragmatic and prudential transgressions, and (in Study 2) playful vocalizations in response to pragmatic transgressions. Study 1 showed that this differential use of vocal tone is used systematically in everyday life. Study 2 allowed us to standardize the context of the maternal intervention and perform additional acoustical analyses. A combination of principal component analysis and linear discriminant analysis applied to pitch and intensity data provided quantitative measures of the differences in vocal responses. These differentiated vocal responses are likely contributors to children's acquisition of norms from early in life.

*Keywords:* vocal expression of emotion, moral development, infancy

*Supplemental materials:* <http://dx.doi.org/10.1037/a0036608.supp>

Vocal emotional signals are well suited to prohibit infants from doing something wrong. When young children violate norms for how to behave, for instance, by hitting a sibling, parents will usually try to stop the child from continuing the transgression. Emotional responses communicated through prosodic vocal features, such as pitch and intensity, are often better suited than other modalities for communicating prohibitions. Unlike physical interventions, vocal expressions do not depend on the child being right next to the parent and, unlike facial expressions, their effect does not depend on the child looking at the parent. Through reliance of prosodic (nonverbal) features of speech, vocal emotional communication can be effective even at ages when children lack advanced linguistic skills. Further suggesting that vocal emotions do play a significant role in parent–infant communication, past studies have

found emotional features to be more prominent in infant-directed speech than in adult-directed speech (Trainor, Austin, & Desjardins, 2000).

The specific emotional tone used to prohibit an infant can contain important information about the mother's perception of the transgression (Dahl, Campos, & Witherington, 2011). The human voice can communicate a variety of qualitatively and quantitatively different emotions (see Juslin & Scherer, 2005). A mother's voice could thus convey whether the mother is angered by, afraid of, or perhaps amused by what the child is doing. When applied in a consistent fashion, such emotional response patterns could help guide the child's acquisition of norms for how to behave, both in the short and long term. Maternal vocalizations could help children understand how their behavior is affecting others or, in the case of fear, how persisting with the action could harm the children themselves ("Am I likely to get hurt if I continue?"). Research within social referencing provides powerful evidence for how specific emotional signals from others can regulate infant behavior, for instance, the avoidance of a dangerous location (Mumme, Fernald, & Herrera, 1996; Sorce, Emde, Campos, & Klinnert, 1985; Vaish & Striano, 2004; Walle & Campos, 2012).

The present research investigated the hypothesis that the nature of infants' transgressions is a key determinant of mothers' emotional tone of voice. We focused on three types of transgressions of particular importance in early childhood: moral transgressions involving harm to others (e.g., hitting), prudential transgressions involving harm to the transgressor (e.g., touching a hot plate), and pragmatic transgressions creating inconvenience (e.g., spilling or playing with breakable family property; Davidson, Turiel, & Black, 1983; Kim, 2013; Smetana, 2013; Turiel, in press). Parents draw conceptual distinctions between these types of transgressions

---

This article was published Online First May 26, 2014.

Audun Dahl, Briana R. Sherlock, Joseph J. Campos, and Frédéric E. Theunissen, Department of Psychology, University of California, Berkeley.

This research was supported by grants from the National Institute of Child Health and Human Development (1R03HD077155-01), Amini Foundation for the Study of Affect, and the Norway American Association, and scholarships from the American Psychological Foundation and the Greater Good Science Center. We thank Elliot Turiel, Sona Patel, Ervin R. Hafter, and Mahesh Srinivasan for helpful suggestions, the research assistants at the UC Berkeley Infant Studies Center for data collection and coding, and the parents and children for their participation.

Correspondence concerning this article should be addressed to Audun Dahl, Institute of Human Development, University of California, Berkeley, 1121 Tolman Hall, MC 1690, Berkeley, CA 94720-1690. E-mail: [dahl@berkeley.edu](mailto:dahl@berkeley.edu)

sions, as do children of preschool age and older. They judge moral and prudential transgressions as more serious than pragmatic transgressions, and give different justifications for why moral, prudential, and pragmatic transgressions are wrong (Kim, 2013; Smetana, Kochanska, & Chuang, 2000; Tisak, 1993). Consistent with the idea that children and adults distinguish between types of transgressions, several studies have found that these three transgression types engender different verbal and behavioral reactions from others (Dahl & Campos, 2013; Smetana, 1989; Tisak, Nucci, & Jankowski, 1996).

Previous studies, however, have not investigated whether moral, prudential, and pragmatic transgressions elicit different emotional vocal reactions from caregivers. Specifically, mothers may respond more angrily when an infant hits someone (a moral transgression) than when a child purposely pours milk on the table (a pragmatic transgression). In contrast, mothers might respond more fearfully when a child is running toward the street (a prudential transgression) than when an infant is engaged in a moral transgression, given that infants are usually not physically strong enough to cause serious physical injury others. These different emotional reactions may be expressed through differing vocal tones (Juslin & Scherer, 2005), for instance, an angry voice, in the case of moral transgressions, and a fearful voice, in the case of prudential transgressions.

The beginning of the second year is a particularly interesting period to study the use of differentiated emotional responses to infant transgressions. During this period, mother–infant interactions are characterized by an increased number of conflicts (Rijt-Plooij & Plooij, 1993). Some conflicts relate to infants' increasing aggression, whereas others pertain to transgressions made possible by improvements in motoric and cognitive abilities, for instance, climbing on furniture (Biringen, Emde, Campos, & Appelbaum, 1995; Hay, 2005). At the same time, infants' linguistic capacities remain limited. By 14 months, 75% of infants are reported to produce less than 50 words and to understand less than 250 words on the MacArthur Communicative Development Inventories (Fenson et al., 1994). Most infants do not produce their first sentence until around the second birthday.

Perhaps because of infants' limited linguistic abilities, mothers rely heavily on nonverbal communication (such as prosodic features of speech) during this period. In general, prosodic aspects of infant-directed speech appear to be more expressive than prosodic aspects of adult-directed speech. Infant-directed speech is characterized by higher pitch, greater pitch variation, and hyperarticulated vowels (Burnham, Kitamura, & Vollmer-Conna, 2002), and is typically perceived as more emotional than adult-directed speech (Singh, Morgan, & Best, 2002; Trainor et al., 2000). These exaggerated prosodic features make it easier for listeners to infer the communicative message in infant-directed speech, for instance, in judging whether an utterance was an attempt to draw attention to something or an expression of approval (Fernald, 1989; Singh et al., 2002; Trainor et al., 2000). For these reasons, we expected that prosodic aspects of mothers' vocal reactions (i.e., their vocal tone) would be especially likely to differentiate between types of transgressions early in the second year.

In two studies, we investigated whether mothers respond with different vocal tones to infants' moral, prudential, and pragmatic transgressions. The following five hypotheses were proposed, based on past research on vocal emotional communication, infant-

directed speech, and parental conceptions of norms, as well as extensive review of video recordings of naturalistic mother–infant interactions.

**Hypothesis 1: Intense angry (stern) vocalizations are associated with moral transgressions.** Several researchers have proposed, and found, that parents place greater emphasis on moral and prudential norms than on pragmatic norms early in the second year (Dahl & Campos, 2013; Gralinski & Kopp, 1993; Smetana et al., 2000). Yet moral transgressions differ from prudential transgressions in several ways. First, there are numerous ways the child can pose a danger to him- or herself (a prudential transgression), but there is a comparably limited and relatively homogeneous class of behaviors that cause harm to others. Parents may therefore begin to think that children should “know better” than to harm others, without, thereby, expecting their children to avoid all prudential transgressions. Next, as noted, moral transgressions at this age usually do not pose serious physical danger to anyone, whereas prudential transgressions can be life threatening for the child and thus more evocative of fear than anger (see Hypothesis 2). In support of the hypothesis that intense anger-like vocalizations would be associated with moral transgressions, Dahl and Campos (2013) found that mothers reported being more angry after moral transgressions than after prudential and pragmatic transgressions. It seems likely that this anger would be, in part, communicated through mothers' vocalizations.

**Hypothesis 2: Fearful vocalizations are associated with prudential transgressions.** According to a common definition, fear is the emotional response to an uncontrollable threat to something of importance to the person (e.g., Lazarus, 1991). Accordingly, fear is the natural response when a mother sees that her infant is about to do something dangerous, such as crawling toward the edge of a tall table. In contrast, the risk of the infant breaking a plastic cup or spilling on the floor (pragmatic transgressions) is presumably seen as less serious and fear-evoking. Correspondingly, mothers report being more afraid when responding to a prudential transgression than when responding to moral or pragmatic transgressions (Dahl & Campos, 2013), supporting our hypothesis that fearful vocalizations will be primarily associated with prudential transgressions.

**Hypothesis 3: Comforting vocalizations are associated with pragmatic and prudential transgressions.** Even when intervening on a transgression, mothers may often be concerned with keeping children from getting upset (Hastings & Grusec, 1998; Ross, 1996). We hypothesized that vocal tones intended to comfort the child during a conflict would be common following pragmatic and prudential transgressions, but not moral transgressions. Pragmatic transgressions are very common in the second year, given infants' interest in motoric and social exploration (Dunn, 1988). For this reason, combined with the fact that the physical consequences of pragmatic transgressions (milk on the table or clothes all over the floor) are generally less serious than those of moral transgressions, we did not expect parents to perceive pragmatic transgressions as very severe. Similarly, many prudential transgressions do not involve an immediate and serious threat to the child. In fact, the most serious dangers to the child's health are often dealt with through prevention, for instance, by removing a dangerous object (Gärling & Gärling, 1995). Moreover, more so than moral transgressions, prudential and pragmatic transgressions constitute a heterogeneous collection of actions and objects, which

likely makes the corresponding norms difficult to grasp for young children. In line with the idea that mothers are less concerned with stressing the wrongness of pragmatic and some prudential transgressions, mothers are more likely to try distract (rather than explicitly prohibit) their child during pragmatic and prudential transgressions than during moral transgressions (Dahl & Campos, 2013). On the basis of these considerations, we expected loving, comforting vocalizations to be especially common in situations involving pragmatic transgressions, but to also occur in some prudential situations. In contrast, we expected comforting vocalizations to be rare in response to moral transgressions.

**Hypothesis 4: Playful vocalizations are associated with pragmatic transgressions.** Pilot observations, as well as conversations with mothers, suggested that pragmatic transgressions sometimes contain game-like or comical elements. For instance, children would play with a prohibited object in order to elicit a reaction from the caregiver and were delighted when they received a prohibition (Dunn, 1988). For this reason, and because pragmatic transgressions were expected to be seen as the least serious, we hypothesized that maternal responses to pragmatic transgressions (but not to moral or prudential transgressions) would sometimes be playful and even include laughter.

**Hypothesis 5: Low-intensity angry (firm) vocalizations are equally frequent after all types of transgressions.** Not all vocalizations could be expected to differentiate among domains. After all, the child is engaging in an unwanted behavior, and a common response from a parent may therefore be a mixture of annoyance and a gentle attempt to remind or inform the child of a norm prohibiting their current behavior. For this reason, we expected that matter-of-fact low-intensity anger ("firmness") vocalizations would be approximately equally common in all three situation types.

The studies described here tested the five hypotheses by investigating mothers' naturally occurring and experimentally elicited responses to infant transgressions. Study 1 assessed whether moral, prudential, and pragmatic transgressions in everyday life elicit different vocal emotional responses from mothers. The laboratory paradigm used in Study 2 allowed us to standardize the context of the maternal vocalizations and perform additional acoustical analyses on the responses.

Before presenting the two studies, a brief discussion of culture is in order. It is commonly reported that parents of European origin in Western cultures express more positive emotion and less negative emotion toward their children than do parents of other cultural and ethnic backgrounds (see Camras, Shuster, & Fraumeni, 2014; Cole & Tan, 2007). Nevertheless, we did not expect to find significant effects of mothers' ethnic background in the present studies.<sup>1</sup> Although there are cultural differences in relative emphasis on certain emotions, there are also large cross-cultural similarities between parents in their responses to infants' behaviors. For instance, parents in most cultures tend to respond negatively to aggressive behaviors (Honig & Chung, 1989; see Cole & Tan, 2007, for a discussion). Insofar as parents consider moral transgressions to be worse than pragmatic transgressions, and prudential transgressions to involve a threat to a person they care about (their child), the predicted differential pattern of emotional responses (Hypotheses 1-5) would hold across different communities. Moreover, our investigations analyzed responses in a relatively small sample of middle-class mothers living in the San

Francisco Bay Area. The region, although ethnically diverse, remains a metropolitan and liberal part of a Western country. Parenting styles in the study population might therefore be relatively homogenous. We will revisit the issue of cultural variability in the General Discussion.

## Study 1: Maternal Responses to Naturally Occurring Transgressions

### Method

**Participants.** Twenty-six families participated in a 2.5-hr home visit when the target child (11 female, 15 male) was 14 months of age ( $M_{\text{age}} = 14.5$  months,  $SD_{\text{age}} = 0.63$ ). The sample consisted predominantly of non-Hispanic Caucasian (70%) and Asian American (23%) upper-middle-class families living in the San Francisco Bay Area.

**Materials and procedure.** In all visits, the target child, the mother, and one older sibling less than 8 years of age were present. All visits included a mealtime. Since the goal of this study was to investigate naturally occurring responses to transgressions in everyday life, mothers were told to ignore the observer and pursue typical daily activities with her children. During the visit, the observer followed the child with a video camcorder (Panasonic HDC-TM900). Every time the mother intervened either by trying to stop the child from doing something or by negatively evaluating something the child had done, the observer logged the occurrence of this event using an iPod Touch (Apple Inc.). Mothers were given a demographics questionnaire to be filled out and returned by mail (25 of 26 mothers returned the questionnaire).

After the visit, the observer watched the logged situations in the video recordings to identify situations in which it was unclear why the mother intervened (for instance, when the mother stopped the child from opening a cabinet but the observer did not know if the cabinet contained something dangerous or something that could create a mess). The observer then conducted a phone interview with the mother about these situations (typically 30% to 40% of the situations). The interviewer gave a brief description of the relevant situation, asked the mother what her concern was in intervening on the child's behavior, and wrote down a summary of the mother's response. The summaries of the mother's response for each situation were used for classifying the child's transgression (see *Coding of video recordings*).

### Coding, data reduction, and data analysis.

**Coding of video recordings.** The video recordings obtained during the home visits were coded by research assistants blind to the study hypotheses. Coders only coded situations that had been logged during the observation. Video coding was done using the software Interact (Mangold, GmbH). Video recordings from 20% of visits were double coded to assess interrater agreement. Agreement was calculated for situations in which both coders had noted at least one intervention from the mother.

In a first pass, coders judged what kind of norm had been violated: *moral* (interpersonal harm), *prudential* (doing something dangerous), *pragmatic* (making a mess, spilling, playing with a

<sup>1</sup> We thank an anonymous reviewer for prompting us to make this clarification.

breakable object, causing annoyance or inconvenience, but no harm), or *other* (not included in the analyses). Only situations in which the mother intervened upon the child's transgression were included. If a mother had been interviewed about the situation, the coder classified the situation based on the mothers' explanation. Situations involving either moral or prudential concerns in addition to pragmatic ones were classified as moral or prudential, respectively. Situations involving both moral and prudential concerns were rare and were excluded from analysis ( $n = 2$ ). Agreement among coders for classifying situations was high, with  $\kappa_{\text{Cohen}} = .90$ .

Coders also assessed how far the child had gotten in the transgressive activity at the time of the first intervention ("degree of completion": *started* or *completed*,  $\kappa_{\text{Cohen}} = .79$ ). For instance, if a child was reaching toward a prohibited object on the table at the time when the mother first intervened, the degree of completion would have been coded as *started*. In contrast, if the child had already pushed the object off the table and onto the floor, the degree of completion would have been coded as *completed*. Degree of completion was coded because we suspected that this could affect the use of certain vocal tones, especially those expressing fear.

**Marking of maternal vocalizations.** The sound from each video file was extracted as a separate Waveform Audio File Format file using the VLC media player (VideoLAN). In a first pass, a coder marked the beginning and end of each maternal utterance in the coded situations. An utterance was defined as a syntactically complete sentence (containing a subject and a predicate, e.g., "that's not nice," or an imperative, e.g., "don't do that"), a prohibitive word occurring outside a complete sentence (e.g., "no," "uh-uh"), or the child's name not followed by a sentence within five seconds (e.g., "John!"). The coder also rated the amount of background noise in the recording on a 5-point scale (1 = *no background noise*; 2 = *some background noise*; 3 = *cross-talk, music, or other noise clearly audible*; 4 = *loud background noise makes most of maternal utterance difficult to hear*; 5 = *speech of mother not discernible or mother is whispering*).

**Coding of vocal tone.** To prevent coders of vocal tone from being informed by the verbal content of maternal vocalizations, a 500-Hz low-pass filter was applied to the audio files (Fernald, 1989; Scherer, 2003) using the open-source software Audacity.

Filtered vocalizations from each mother were coded three times, as the use of multiple coding passes has been found to improve the

reliability of emotion coding of vocalizations (Shrivastav, Sapienza, & Nandur, 2005). In the present case, pilot coding suggested that using more than three passes would not further increase interrater agreement. In each coding pass, the vocalizations were presented to the coder in random order, preventing the coder from taking the sequence of utterances in a given situation into account when coding vocal emotion. The coder was also unaware of the type of situation to which the mother was responding.

Table 1 lists the coded vocal tones, their definitional features, and the situation type(s) with which we expected the tones to be associated. Each vocal tone was given a name by combining the low- and high-intensity variants of the relevant dimension (for instance, "firm" is the low-intensity end and "stern" is the high-intensity end of the tone labeled *Firm-Stern* [FS]). Definitions for the four codes were developed from findings in past studies of vocal emotion in infant-directed speech (Fernald, 1989; Trainor et al., 2000) as well as review of pilot data. FS was an "angry" tone characterized by stable, low pitch. High-intensity variants of the FS tone had high mean amplitude and abrupt changes in amplitude. *Worried-Scared* (WS) was a "fearful" tone defined by generally high pitch, quick pitch changes, short word duration, and high amplitude. *Warm-Comforting* (WC) was a "loving" tone characterized by large but slow variability in pitch, drawn-out vowels, and low amplitude. Finally, *Playful-Laughing* (PL) was a "joyful" tone with quick and large changes in pitch and low amplitude that, unlike the other vocal tones, could contain laughter. The online supplementary materials include audio clips and spectrograms for five maternal vocalizations illustrating high-intensity FS (stern), low-intensity FS (firm), WS, WC, and PL tones.

In each coding pass, the coder assigned one of the codes listed in Table 1 as the primary vocal category for a given utterance. The coder also indicated intensity for this code on a scale from 1 to 5. If the coder found that a second tone was also present in the utterance, she had the option of assigning a secondary vocal category to the utterance. In cases for which two tones appeared to be equally present, the tone present toward the end of the utterance was considered primary.

**Data reduction.** Codes from the three coding passes were combined in the following way: If a vocalization had been coded as WS two or more times (i.e., two times out of the three to six primary plus secondary codes assigned to a given clip), the vocalization was classified as WS. This rule was implemented because it was found that WS tones very often co-occurred with other

Table 1  
Vocal Categories: Definitional Features

Acronym (High-Low)	Acoustical property								
	Mean pitch	Pitch Range	Pitch change	Word duration	Speech speed	Mean amplitude	Amplitude variability	Can contain laughter	Hypothesized situation type
FS (Firm-Stern)	Low	Narrow	Int.med.	Int.med.	Int.med.	High (+)	High (+)	No	Moral
PL (Playful-Laughing)	High	Wide	Int.med.	Int.med.	Int.med.	Low	Low	Yes	Pragmatic
WC (Warm-Comfort)	Int.med.	Wide	Slow	Long	Slow	Low	Low	No	Pragmatic (Prudential)
WS (Worried-Scared)	High	Wide	Quick	Short	Fast	High	Int.med.	No	Prudential

*Note.* The table lists the four vocal coding categories and their definitional features used by coders in each coding pass. Coding data from the three passes were combined according to the data reduction diagram in Figure 1 (see Study 1, Method). A "+" sign indicates that the property is associated only with the high-intensity variant of the category (e.g. FS-high). Int.med = intermediate.

tones, especially FS. Otherwise, the vocalization was classified according to whichever tone had been coded most frequently (primary or secondary). If two codes had been used equally often, whichever code had been used more often as *primary* was taken as the final classification (if the number of primary codes was also identical, the vocalization was classified as "mixed"). If a vocalization was classified as FS and had a mean intensity rating of 3 or greater, the vocalization was classified as FS-high; otherwise, it was classified as FS-low. This data reduction scheme is shown in Figure 1.

**Reliability.** Data from 20% of participants was double coded to assess interrater agreement. Ninety-four percent of clips receiving a noise rating of less than 4 (i.e., clips in which the mother's voice was clearly audible in the filtered recordings) were classified as falling into one of the five emotion categories (i.e., not mixed). Agreement for classifying vocalizations as FS-low, FS-high, WS, WC, or PL was  $\kappa_{\text{Cohen}} = .73$ . It was found, however, that it was particularly difficult to identify PL vocalizations in the low-pass filtered recordings. Not only was agreement relatively low, but many of the vocalizations identified as PL in the filtered recordings did not sound like PL to the same coder when listening to the unfiltered recording. The PL category was therefore excluded from the analyses in Study 1. Agreement for the four categories (FS-low, FS-high, WS, WC) used in the analyses was  $\kappa_{\text{Cohen}} = .81$ .

**Data analytic strategy.** There was large variability in the number of maternal interventions per situation, ranging from 1 to 31. Only interventions up to and including the ninth maternal intervention were used in the analyses (the criterion was chosen so as to include 90% of all maternal vocalizations).

As an overall test of whether maternal tone of voice depended on transgression type, we conducted a Pearson chi-square test on the 3 (situation: moral, prudential, pragmatic)  $\times$  5 (tone: FS-low,

FS-high, WS, WC, other) contingency table. Note that this analysis did not take into account the nonindependence of data obtained from the same participant. The data were subsequently analyzed separately for each vocal category using generalized linear mixed models (GLMMs) with binomial error distribution and logistic link function (Hox, 2010). The GLMMs included a random intercept for subjects and a fixed effect of situation type (see Appendix). Level of completion was included in the final model when the respective effect was significant. Preliminary statistical analyses revealed no significant effects of mothers' ethnicity (Caucasian, Asian American, or Other); hence, this variable was not included in the reported models. Hypotheses were tested using likelihood ratio tests by comparing the model deviance difference ( $D$ ) between full and restricted models to a central chi-squared distribution with degrees of freedom equal to the difference in the number of parameters (Hox, 2010). Models were estimated and tested in R (version 2.15.2).

## Results

In 608 situations, mothers were coded as having at least one verbal intervention. Of these situations, 28 were coded as moral, 235 as prudential, and 345 as pragmatic. This means that mothers intervened vocally on children's transgressions approximately nine times per hour. Ninety-seven percent of vocalizations fell into the categories listed in Figure 1 (FS-high, FS-low, WS, WC, or PL).

**Overall chi-square test.** The chi-square test revealed a significant association between situation type and vocal tone,  $\chi^2(8, N = 1665) = 21.16, p = .007$ . Next we report the results from the GLMM analyses.

**FS-high.** There was a significant effect of situation type on the use of FS-high vocalizations,  $D(2) = 6.17, p = .046$  (see Figure 2).

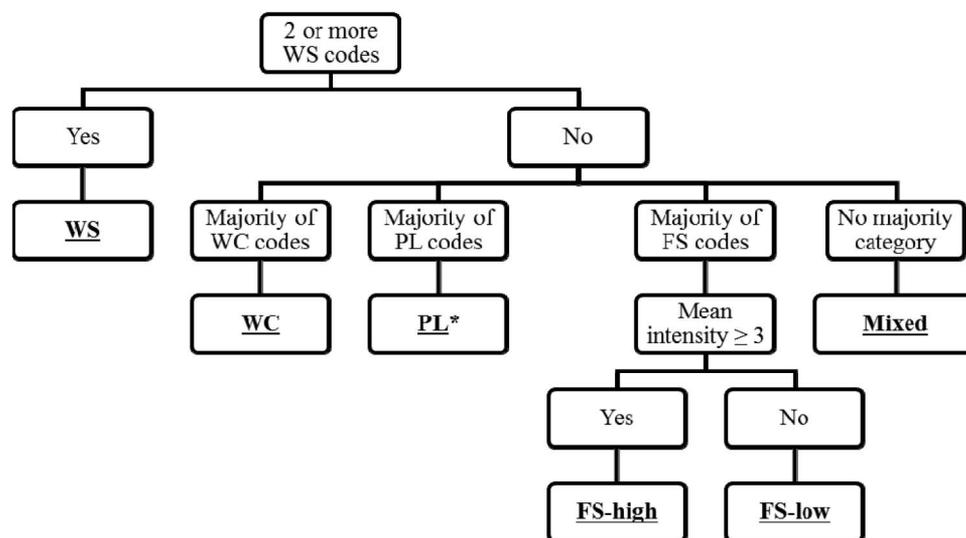


Figure 1. Data reduction scheme. The diagram shows how coding data from the three coding passes were combined for each vocalization. When deciding which code was used most often, both primary and secondary codes were counted. If two codes were equally frequent, whichever code had been used most often as primary was used taken as the majority code. The boldfaced and underlined code shows the final code assigned to each vocalization. WS = worried-scared; WC = warm-comforting; PL = playful-laughing; FS = firm-stern. \* PL vocalizations were only analyzed in Study 2 (see Study 1 methods).

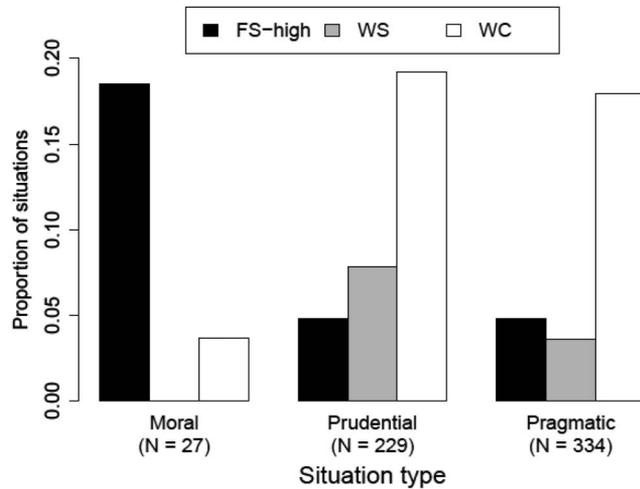


Figure 2. Vocal tone as a function of situation type (Study 1). Bars show the proportion of situations of each type in which mothers used the respective vocal tones at least once. Numbers in parentheses show number of situations in each category.

The use of FS-high vocalizations was higher in the moral situations (19% of situations) than in prudential situations (5%),  $D(1) = 5.57, p = .018$ , and pragmatic situations (5%),  $D(1) = 5.89, p = .015$ . There was no effect of level of completion,  $D(1) = 0.77, p = .78$ , on the use of FS-high vocalizations.

**FS-low.** The use of FS-low vocalizations did not depend significantly on situation type,  $D(2) = 1.38, p = .50$ . There was also no effect of level of completion on the use of FS-low vocalizations,  $D(1) = 1.06, p = .30$ . Across situation types, 80% of situations contained at least one FS-low vocalization.

**WS.** The use of WS vocalizations also depended on situation type,  $D(2) = 7.38, p = .025$  (see Figure 2). As predicted, WS vocalizations were the most common in prudential situations. Further analyses revealed that the difference between prudential (8%) and moral situations (0%) was significant,  $D(1) = 4.64, p = .031$ . In contrast, the difference between prudential and pragmatic situations (4%) did not reach statistical significance,  $D(1) = 3.34, p = .067$ , nor did the difference between moral and pragmatic situations,  $D(1) = 3.21, p = .073$ . The use of WS vocalizations were significantly more common in situations in which the child had completed at least one transgressive act (12%, vs. 4%),  $D(1) = 5.57, p = .018$ .

**WC.** WC vocalizations also depended significantly on situation type,  $D(2) = 6.86, p = .032$  (see Figure 2). WC vocalizations were more common in prudential (19%) than moral situations (4%),  $D(1) = 5.39, p = .020$ . They were also more common in pragmatic situations (18%) than in moral situations,  $D(1) = 6.19, p = .013$ . In contrast, there was no difference between prudential and pragmatic situations,  $D(1) = 0.38, p = .54$ . There was also no effect of level of completion on the use of WC vocalizations,  $D(1) = 0.09, p = .76$ .

## Discussion

The naturalistic data show that mothers' vocal interventions on transgressions occur frequently in the everyday life of young

children, consistent with past reports (Dunn, 1988; Kuczynski, Kochanska, Radke-Yarrow, & Girmius-Brown, 1987). In the present study, we recorded an average of nine intervention episodes per hour, which is similar to rates found in other studies of caregiver-infant interactions (Power & Parke, 1986). The prominence of these situations in infants' lives means they could have a major impact on early social and cognitive development.

As predicted, mothers' vocal emotional responses to child transgressions depended on the nature of the child's transgression. Mothers were more likely to respond to moral transgressions with intense FS vocalizations, more likely to respond to prudential transgressions with WS vocalizations, and more likely to respond to pragmatic and prudential transgressions with WC responses.

These data show that prosodic features of maternal interventions on child transgressions contain information about the nature of the violated norm. These emotional signals may help children construct an understanding of the difference in meaning and importance between moral, prudential, and pragmatic norms.

Study 1 suffered from a few limitations. First, differences between other, nonvocal forms of maternal intervention would add noise to the relation between situation type and vocal response. Dahl and Campos (2013) found that interventions on moral, prudential, and pragmatic transgressions also differ, for instance, in the use of physical intervention and distraction. To rule out these potential confounds, it was therefore desirable to see whether mothers' vocal responses would also differ by transgression type if mothers could *only* respond vocally. Second, the number of moral situations was quite low, limiting statistical power and making parameter estimates less reliable. Third, Study 1 could not assess differences in the tendency *not* to respond at all to certain transgressions. For instance, if mothers consider pragmatic transgressions less important than moral transgressions, mothers would presumably be less likely to intervene on pragmatic transgressions than on moral transgressions. Finally, the variability in content, duration, and recording quality between the maternal utterances made it difficult to use quantitative and objective acoustical measures of vocal characteristics (such as pitch and intensity; Scherer, 2003). In Study 1, we were therefore forced to rely on the qualitative categorizations obtained from human coders.

Study 2 addressed these limitations by experimentally eliciting prohibitive vocalizations from mothers. In this study, we showed mothers clips of infant transgressions and asked them to pretend they were telling the child not to do what the child was doing. Eliciting maternal vocalizations in the laboratory allowed us to improve recording quality by having the mother speak into a microphone and by eliminating background noise. In Study 2, we were also able to standardize the verbal content of the vocalizations by instructing the mothers always to use the same phrase, which made the prosodic features of the vocalizations more easily comparable. Furthermore, the standardization of the verbal content meant that the coders could listen to the unfiltered vocal recordings without the verbal content revealing the kind of transgression to which the mother was responding.

The improved sound quality and the standardization of verbal content in Study 2 allowed us to analyze the pitch and intensity data from the vocalizations. Using quantitative analysis of these acoustical measures, in addition to the analyses of the human coding, had several advantages. Most importantly, it allowed us to classify the mothers' vocalizations without relying on a priori

vocal categories based on acoustic features salient to an adult listener. Thus, it allowed us to assess whether a learner could perceive differences between vocalizations elicited by moral, prudential, and pragmatic transgressions without already possessing vocal categories like “stern” or “fearful.”

To accomplish this, we used a combination of principal component analysis (PCA) and linear discriminant analysis (LDA; Theunissen, Mouterde, & Mathevon, 2013). By selecting a subset of the principal components (PCs) for the time-varying pitch and intensity profiles of each utterance, we could reduce the dimensionality of the data without losing large amounts of information (as opposed to, e.g., using the mean pitch or intensity across time). The PCs were, in turn, used in LDA models to predict what kind of transgression had elicited the vocalizations. To our knowledge, this statistical procedure has never been used in the analysis of human emotional vocal communication. On a methodological level, Study 2 thus introduces a potentially useful tool for addressing a number of questions in the study of human vocal expression.

## Study 2: Experimental Elicitation of Maternal Vocalizations

### Method

Thirty-five mothers of 14-month-old infants (16 female, 19 male;  $M_{\text{age}} = 14.5$  months,  $SD_{\text{age}} = 0.54$  months) participated in this study. Fifty-seven percent of mothers were Caucasian, 11% were Asian, 11% were African American, and 9% were Hispanic. Twenty-nine percent had a graduate degree and 40% had a college degree. The mean age of mothers was 33.7 years ( $SD = 6.54$ ).

#### Materials.

**Experimental stimuli.** The experimental stimuli in Study 2 were video recordings of child transgressions obtained in Study 1. If a mother in Study 1 had given written permission to use the obtained video recordings in other experiments, research assistants naïve to the study hypotheses viewed the moral, prudential, and pragmatic situations coded in that visit. Research assistants assessed whether the recorded situation was suitable as a stimuli in Study 2 based on the following criteria: The video segment had to contain a situation in which (a) the child was clearly visible immediately prior to the transgression; (b) the transgression was detectable given only the information present in the clip (for instance, a transgression in which a child approached a prohibited cabinet would not be suitable because one would not know what was in the cabinet simply from watching the video recording); and (c) the clip could be edited so that no interventions from other family members would be visible or audible.

Three sets of video clips were created, each containing 12 clips: four moral, four prudential, and four pragmatic clips. Clips lasted 2 to 5 s. At the end of each clip, the final image frame was frozen for another 3 s, giving mothers additional time to respond. For each video set there were also six extra clips (two moral, two prudential, and two pragmatic), to be used if a mother chose not to respond to ordinary clips falling in the corresponding category (see the following Procedures section). The order of the 12 ordinary clips within each set was random, with the constraint that a mother would never watch two clips in the same category in a row.

**Recording set-up.** The mothers' vocalizations were recorded using a Shure SM57 unidirectional dynamic microphone, an EMU

0204 External Sound Card connected to a desktop computer, and the software Audacity. The microphone was positioned approximately 6 in. from the mothers' mouth and behind a pop-screen.

**Questionnaires.** Two questionnaires were administered: a demographics questionnaire and a behavior rating checklist. On the behavior rating questionnaire, mothers indicated how important it was for them to encourage their child to refrain from or engage in a set of behaviors on a scale from 1 (*unimportant/never requested*) to 4 (*very important*). The list of behaviors was taken from Gralinski and Kopp (1993). For the purposes of the present study, only ratings of behaviors involving violations of moral norms (not being rough with other children, not biting mom or dad), prudential norms (not climbing on furniture, not going into the street, not touching things that are dangerous), and pragmatic norms (not coloring on walls, not tearing up books, not interrupting, not playing with food, not spilling drinks) were included. For each mother, mean importance ratings for moral, prudential, and pragmatic transgressions were calculated.

**Procedures.** Mothers were told that they would watch some video clips of infants doing various things that could be considered wrong. They were told to respond vocally to these clips as if they were telling the child not to do whatever the child was doing. The researcher emphasized that the mother should *not* vocalize if the clip showed a behavior upon which they would not normally intervene. Importantly, mothers were not informed about the study hypotheses or about the three categories of transgressions shown in the video clips (moral, prudential, pragmatic). They were told to use the following standardized phrase when responding: “No, don't do that.” The phrase was shown visually to mothers so as not to lead mothers to say the phrase in a particular way. Mothers could respond at any point during the video clip.

The recording session began by calibrating the microphone gain to avoid clipping effects. Mothers were asked to say the phrase one or more times and the experimenter adjusted the microphone gain so that the maximum amplitude of the vocalization was approximately 12 dB below the maximum recording amplitude in Audacity.

Mothers were randomly assigned to view one of the three video sets. Mothers first watched the 12 ordinary clips in the assigned set. The playback of video clips was controlled by an experimenter, who ensured that the mother was ready before playing back each clip. If a mother had failed to respond to one or more of the 12 ordinary clips, she was then shown extra clips in the appropriate category (up to two clips in each category). For instance, if a mother failed to respond to one prudential clip and two pragmatic clips among the ordinary clips, she would watch one prudential and two pragmatic extra clips. All mothers thus watched between 12 and 18 clips total.

**Coding, data reduction, and reliability.** All audio clips were coded without knowledge of the type of transgression to which the mother responded. (Because mothers always used the same phrase, regardless of transgression type, low-pass filtering was not necessary in Study 2.) The coding and data reduction procedures in Study 2 were identical to those in Study 1 (see Table 1), except that the coder was also allowed to indicate intensity for the secondary vocal code (if a secondary code was used), and that the coder would mark “no response” if the mother chose not to respond to a particular video clip.

Table 2 shows mean values of the definitional acoustical features for each of these categories, which were consistent with definitions given in Table 1. Figure 3 shows example plots for pitch and intensity as a function of time for each of the five final vocalization categories (FS-high, FS-low, WS, WC, and PL).

Data from 20% of participants were double coded to assess interrater agreement. Ninety-six percent of vocalizations were classified as falling into one of the emotion categories by both coders. Agreement for classification of vocalizations as FS-low, FS-high, WS, WC, or PL was  $\kappa_{\text{Cohen}} = .74$ . As the coded recordings were not low-pass filtered, vocalizations classified as PL were reliably coded and therefore included in Study 2.

In order to analyze pitch and intensity, a research assistant manually removed all silence at the beginning and end of each vocalization recording. The edited Waveform Audio File Format files were then imported into the software Praat (ver. 5.3.52). Pitch and intensity were estimated in Praat using the autocorrelation method, with window length set to 10 ms, allowing for capturing of pitch and amplitude modulations of up to 50Hz. The pitch estimation was checked visually by a research assistant, using the Praat plotting functions, to see if pitch had been estimated for periods of silence or if the pitch estimate deviated substantially from the fundamental frequency. Estimation errors were marked and the estimated pitch was extracted by removal (in the case of silence) or interpolation (in the case of wrongly estimated pitch, using 10 samples before and 10 samples after the period for which pitch had been misestimated). The research assistant also marked the beginning and end of the word “no” in the maternal utterances and noted whether the mother said the correct phrase or had included additional words (e.g., “No, no, no, don’t do that”) or omitted words (e.g., “Don’t do that”). Only recordings in which the mother used the correct phrase were analyzed. Pitch and intensity data were analyzed in R. Intensity estimates for each participant were mean-centered before analyses.

#### Data analytic strategy.

**Human coding.** Because we had specific hypotheses about which vocalization types would and would not be affected by transgressions type, separate analyses were conducted for each coded response types (FS-high, FS-low, WS, WC, PL, and no response). Similar to the procedures in Study 1, we used both a

Pearson chi-square test on the Tone  $\times$  Situation Type contingency table and logistic GLMMs fitted separately for each response type to take into account the within-subject structure of the data (see Appendix for sample equations). As in Study 1, statistical analyses revealed no significant effects of mothers’ ethnicity; hence, this variable was not included in the reported models.

**Importance ratings.** Importance ratings were analyzed using a Friedman rank sum test and Wilcoxon signed-ranks test (for pairwise comparisons), as the rating data were negatively skewed and discontinuous.

**Sound analysis.** The main purpose of the analyses of the pitch and intensity data was to predict the classifications of the mothers’ vocalizations without relying on a priori categories (such as the vocal tone categories used by the human coder). We wanted to use this procedure to predict what kind of transgression type (video clip) had elicited the mother’s vocalization.

One suitable and relatively straightforward procedure is that used by Theunissen and his colleagues (2013), in which a subset of PCs of the raw data are used as predictors in a LDA. The raw time-varying pitch and intensity profiles have too many correlated dimensions (corresponding to a data point every 10 ms for each utterance) to be used as predictors in LDA models, unless the number of sound samples is extremely large. PCA allows for reduction of the dimensionality of such data sets without effective loss of the information bearing variability. A subset of the PCs (those with the highest eigenvalues) can then be used as predictors in LDA models, predicting, for instance, what kind of transgression type elicited the vocalization.

To further standardize our vocalization data set for these analyses, we restricted our analysis to the “no” part of the mothers’ vocalizations. The duration of these “no” vocalizations ranged from 0.14 to 1.09 seconds. Visual inspection of the distribution of durations revealed no outliers. Pitch and intensity after the “no” had ended were set to zero.

Next, PCA was conducted on the pitch and intensity data. Our data reduction approach involved using the first  $m$  PCs for pitch and the first  $n$  PCs for intensity in an LDA model predicting situation type (moral, prudential, or pragmatic) or vocal tone assigned by the human coder. To determine how many PCs to

Table 2  
Vocal Categories: Mean Acoustic Properties (Study 2)

Acronym	Acoustical property						
	Mean pitch (Hz)	Pitch range (Hz)	Quickness of pitch change	Word duration (sec)	Utterance duration (sec)	Mean amplitude (dBcent.)	Mean SD of amplitude
FS-high	223.2	153.0	344.8	0.98	1.29	<b>2.49</b>	<b>5.69</b>
FS-low	207.8	127.2	267.9	0.90	1.11	-0.30	4.37
PL	230.1	162.0	337.5	1.02	1.24	-1.42	3.90
WC	224.6	<b>172.1</b>	266.5	<b>1.43</b>	<b>1.63</b>	-0.99	4.30
WS	<b>264.1</b>	<b>171.7</b>	<b>403.9</b>	0.82	1.03	<b>2.17</b>	4.84

*Note.* The table lists the vocal categories used in the analyses. Boldface indicates that the category was defined as *high* on the particular dimension, and italic indicates that the category was defined as *low*. A measure of quickness of pitch change was obtained by regressing pitch onto time separately for each 100-ms segment (using ordinary least squares regression) and taking the average of the absolute value of the slope coefficient. A higher average absolute value of the slope coefficient indicates that the pitch changed more quickly with time. Word duration is the mean total time the participant was vocalizing (excluding breaks). Utterance duration is the total time of the utterance (including breaks), which is an indirect measure of speech speed. FS = Firm-Stern; PL = Playful-Laughing; WC = Warm-Comforting; WS = Worried-Scared.

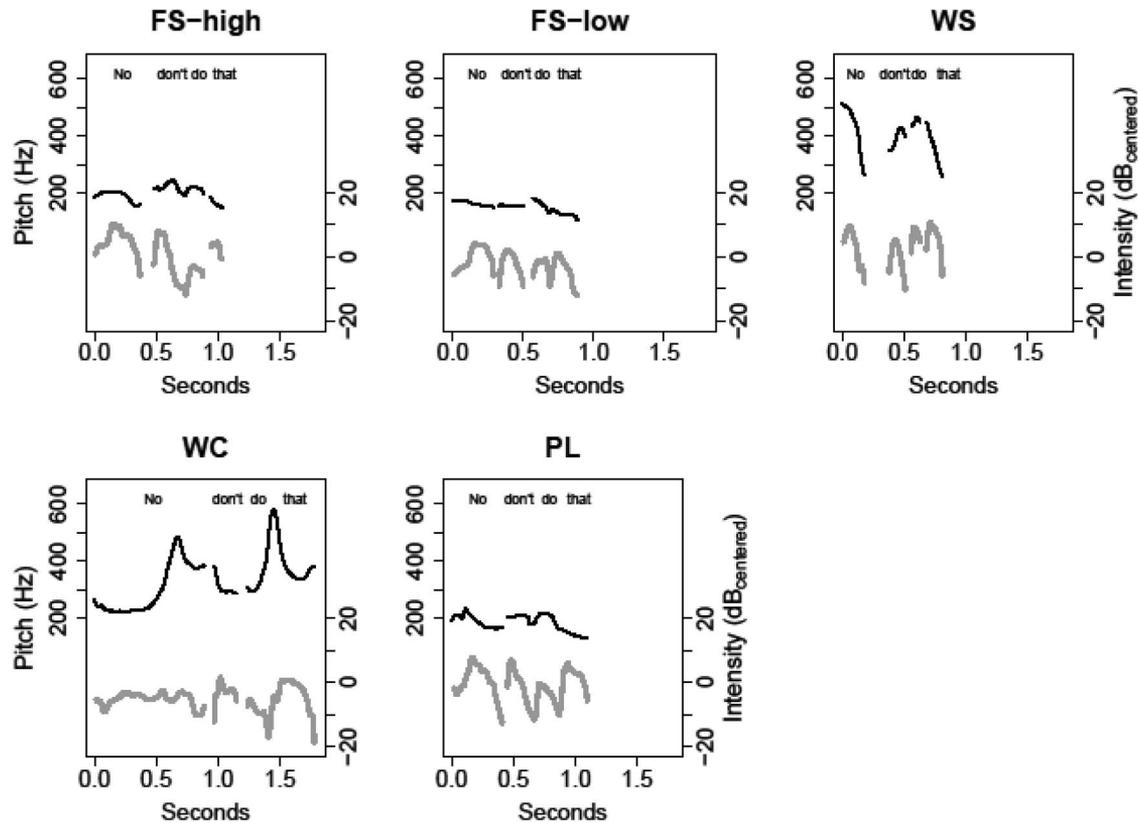


Figure 3. Example plots of pitch and intensity for the final vocal categories (Study 2). The plots show pitch (black line) and intensity (gray line) as a function of time for five vocal recordings. The title of each plot is the final code assigned to each of the clips. See the online supplementary materials for the corresponding audio recordings and spectrograms.

include in the final analyses (the values of  $m$  and  $n$ ), we tried the entire range of 5 to 50 PCs for pitch and intensity, and used LDA to predict situation type and assessed predictive accuracy by cross-validation. For each of 40 iterations of the cross-validation procedure, we randomly selected 95% of clips as a training set and used the remaining 5% as a test set. Performance of the LDAs was evaluated by looking at the proportion of clips within each situation type that were classified correctly. The best model was defined as the model with the highest average *minimum* proportion of correctly classified clips. Average minimum proportion was chosen as a criterion to ensure that the model performed reasonably well for all three situation types.

The best performance was found for  $m = 25$  (pitch PCs) and  $n = 9$  (intensity PCs). The coefficients of this reduced number of PCs were then used in the final LDA models. In other words, 25 values were used to describe the time-varying pitch profile and 9 values were used to describe the time-varying intensity profile of each vocalization. The LDA models used 34 acoustical parameters to predict the situation type. For evaluation of the final models, the cross-validation procedure was repeated 100 times. In addition, a multinomial logistic regression of situation type using the coefficients from the 34 PCs was conducted to provide a test of statistical significance (likelihood ratio test).

## Results

**Human coding.** Mothers were shown a total of 538 video clips of child transgressions. Mothers chose to respond to 338 of these clips: 136 moral clips, 113 prudential clips, and 89 pragmatic clips. Two participants did not respond to any pragmatic clips; the data for these participants were therefore excluded from the analyses. Among all vocalizations, 98% were classified as falling into one of the vocal categories of interest (i.e., FS-high, FS-low, WS, WC, or PL, but not “mixed”).

**No response.** The probability of not responding depended significantly on the type of transgression showed in the clip,  $D(2) = 72.77, p < .001$ . Mothers were more likely to respond to moral clips (85%) than to prudential clips (62%),  $D(1) = 27.92, p < .001$ , and pragmatic clips (46%),  $D(1) = 72.53, p < .001$ . Mothers were also more likely to respond to prudential clips than to pragmatic clips,  $D(1) = 11.67, p < .001$ .

**Overall chi-square test.** Among the clips responded to, there was a significant association between situation type and vocal tone,  $\chi^2(10, N = 329) = 47.06, p < .001$ .

**High FS (FS-high).** As predicted, there was a significant situation effect on the use of FS-high vocalizations,  $D(2) = 24.62, p < .001$  (see Figure 4). Mothers were more likely to respond with FS-high vocalizations in response to moral clips (21%) than to

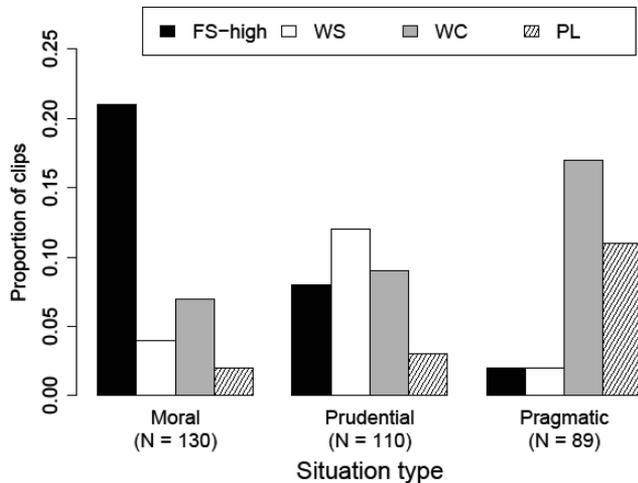


Figure 4. Vocal tone as a function of situation type (Study 2). Bars show the proportion of situations of each type in which mothers used the respective vocal tones at least once. Numbers in parenthesis indicate number of vocalizations for each situation (clip) type.

prudential clips (8%),  $D(1) = 8.54, p = .003$ , and pragmatic clips (2%),  $D(1) = 24.11, p < .001$ . Prudential and pragmatic clips did not differ significantly,  $D(1) = 3.64, p = .056$ .

**Low FS (FS-low).** Consistent with predictions, there was no effect of transgression type on the use of FS-low,  $D(2) = 0.46, p = .79$ . Overall, FS-low vocalizations were used in response to 65% of clips.

**WS.** The use of WS also depended significantly on transgression type,  $D(2) = 9.41, p = .009$  (see Figure 4). WS vocalizations were more commonly used in response to prudential clips (12%) than moral clips (4%),  $D(1) = 5.87, p = .015$ , and pragmatic clips (2%),  $D(1) = 7.27, p = .007$ . In contrast, moral and pragmatic clips did not differ in their tendency to elicit WS vocalizations,  $D(1) = 0.46, p = .50$ .

**WC.** The probability of WC vocalizations also depended significantly on transgression type,  $D(2) = 7.11, p = .029$  (see Figure 4). WC vocalizations were more common in response to pragmatic transgressions (17%) than in response to moral transgressions (7%),  $D(1) = 6.46, p = .011$ . The use of WC vocalizations in response to pragmatic and prudential clips (9%) did not differ significantly,  $D(1) = 3.11, p = .078$ , nor did moral and prudential clips differ significantly in their propensity to elicit WC vocalizations,  $D = 0.44, p = .51$ .

**PL.** Finally, there was also a significant effect of transgression type on PL vocalizations,  $D(2) = 11.20, p = .004$  (see Figure 4). The use of PL vocalizations was more common in response to pragmatic transgressions (11%) than in response to moral transgressions (2%),  $D(1) = 10.02, p = .002$ , and prudential transgressions (3%),  $p = 5.93, p = .015$ . In contrast, moral and prudential clips did not differ in their tendency to elicit PL vocalizations,  $D = 0.46, p = .50$ .

**Importance rating.** There was a significant effect of transgression type on the importance rating, Friedman rank sum test,  $\chi^2(2) = 46.03, p < .001$ . Follow-up analyses revealed that mothers rated it significantly more important to prevent infants from performing moral transgressions ( $M_{importance} = 3.51, SD = 0.68$ ) than

prudential transgressions ( $M_{importance} = 3.18, SD = 0.67$ ), Wilcoxon's test  $V = 401.5, p = .003$ , and pragmatic transgressions ( $M_{importance} = 2.20, SD = 0.72$ ),  $V = 588, p < .001$ . In turn, requesting compliance with prudential norms was rated as more important than requesting compliance with pragmatic norms,  $V = 579, p < .001$ .

Mothers who gave higher average importance ratings for intervening moral behaviors were more likely to respond verbally to moral clips,  $D(1) = 4.51, p = .034$ , and higher ratings for prudential behaviors predicted greater likelihood of responding to prudential clips,  $D(1) = 5.58, p = .018$ . In contrast, the relation between rating of pragmatic behaviors and tendency to respond to pragmatic video clips was not significant,  $D(1) = 1.95, p = .16$ .

**Acoustical analysis.** Mothers used the appropriate verbal phrase ("No, don't do that") in 292 out of the 338 recorded vocalizations. Excluding data for the two mothers who did not respond to any pragmatic clips left 285 vocal recordings for the analyses reported here. Table 3 shows the mean and coefficient of variation (standard deviation divided by the mean) for the pitch and intensity of the included vocalizations.

Table 4 shows the summary of the cross-validation of the LDA models predicting which situation type had elicited a vocalization from the 25 pitch PCs and the nine intensity PCs. Vocalizations elicited by the three situation types were quite well discriminated. On average, 53% of clips from moral situations, 54% of clips from prudential situations, and 46% of clips from pragmatic situations were correctly classified, which is appreciably higher than chance (33%). We also fitted a multinomial logit model predicting situation type from the same PCs used in the LDA. As expected, this model showed a significant effect of the PCs on situation type, likelihood ratio test,  $D(68) = 127.91, p < .001$ .

To further interpret the LDA, we first did a separate LDA including all the data predicting situation type. The main plot in Figure 5 shows the distribution of clips within each situation category as a function of the two linear discriminants. The figure shows that the first discriminant tends to discriminate moral from pragmatic clips, whereas the second discriminant tends to differentiate prudential from moral and pragmatic clips. The figure also indicates the central tendency (median) of each emotional category on the two linear discriminants.

The margin plots in Figure 5 show the results of multiplying the matrix of eigenvectors for the 25 pitch PCs and the nine intensity PCs with the vectors of corresponding scaling values for the two linear discriminants obtained in the LDA (this was done separately for pitch and intensity to facilitate interpretation). Large values on the first discriminant indicates a pitch pattern that was generally stable and peaking between 0.1 to 0.4 s, as well as high and variable intensity (especially 0.1 to 0.4 s), consistent with the FS-high vocal tone. Large values on the second discriminant indicates generally high pitch (especially in the first 0.1 s), highly variable pitch (especially 0.1 to 0.5 s), high intensity in the first 0.2 s, and then highly variable intensity subsequently (0.4 to 0.6 seconds), consistent with the WS vocal tone. Low values on both discriminants (low, stable pitch plus low, stable intensity) were consistent with the WC tone. Indeed, the LDA performed especially well in classifying recordings coded by the human coder as FS-high (58% correct) and WC (46% correct). In contrast, the rates of correct situation classification of WS clips (35%) and PL clips (31%) were close

Table 3  
Study 2: Descriptive Statistics For Pitch and Intensity

Feature	Statistic	Situation type					
		Moral		Prudential		Pragmatic	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pitch (Hz)	Mean	215.57	37.34	219.50	44.61	210.50	35.71
	CV	0.17	0.06	0.15	0.05	0.17	0.08
Intensity (dB <sub>centered</sub> )	Mean	0.13	3.49	0.91	3.88	-1.03	3.15
	CV	0.07	0.02	0.06	0.02	0.06	0.01

Note. CV = coefficient of variation (ratio of sample standard deviation and sample mean).

to chance (33%). Interestingly, the LDA “outperformed” the human coder on the FS-low clips. Even though the FS-low tone was about equally frequent in response to all three transgression types, the LDA model classified 51% of these clips correctly.

## Discussion

Study 2 corroborated the findings from Study 1: Intense angry (FS-high) vocalizations were most common in response to moral transgressions, fearful (WS) vocalizations were most common in response to prudential transgressions, and comforting (WC) vocalizations were more common in response to pragmatic transgressions than in response to moral transgressions (with responses to prudential transgressions falling in between). Data from Study 2 also supported the prediction that PL vocalizations would be the most common in response to pragmatic transgressions.

We also found that mothers were more likely to respond to moral and prudential transgressions than to pragmatic transgressions, consistent with the proposition that moral and prudential norms are more important to mothers than pragmatic norms (Dahl & Campos, 2013; Gralinski & Kopp, 1993). Indeed, in the present study, mothers rated interventions on moral and prudential transgressions as more important than interventions on pragmatic transgressions. On an individual level, mothers' importance ratings of moral and prudential norms significantly predicted whether they would respond vocally to clips showing violations of such norms.

The improved recording quality in Study 2 allowed for analysis of acoustical measures of the mothers' vocalizations. When

using a combination of PCA and LDA, we found that vocalizations in response to moral, prudential, and pragmatic transgressions tended to differ in pitch and intensity patterns. As with the human coding, there was not a complete separation between vocalizations elicited by the three clip types. The relation between the child's transgression and particular maternal responses appears to be probabilistic rather than deterministic (Dahl & Campos, 2013). Some of the variability in responding may be due to stable differences between mothers, and some is likely also due to other situational features, for instance, a given mother's emotional and cognitive state at the time of intervention.

## General Discussion

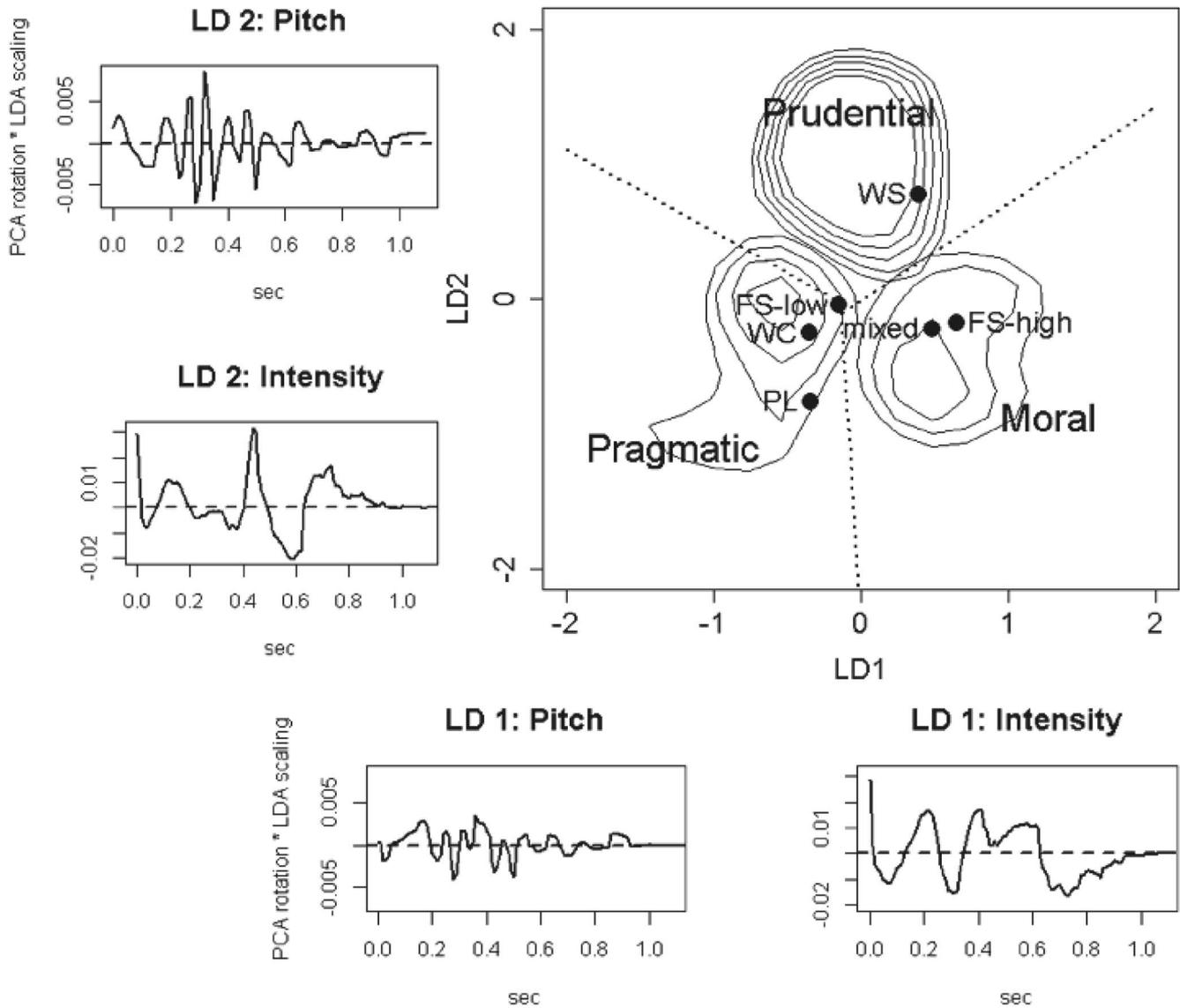
The present studies show that mothers' vocal emotional responses to infant transgressions constitute rich and commonplace features of infants' everyday life. In both studies, mothers tended to respond with different tones of voice to different types of transgressions. As predicted, intense anger-like tones (stable pitch, high mean and large variability of intensity) were associated with moral transgressions, fearful tones (high mean and large variability of pitch, high intensity) were associated with prudential transgressions, and comforting tones (low mean pitch and slow pitch changes, low intensity) were associated with pragmatic (and to some extent prudential) transgressions. In addition, Study 2 found that playful tones are most commonly used in response to pragmatic transgressions, and that pragmatic transgressions are more likely than moral and prudential transgressions to elicit no response at all. Study 2 corroborated the findings from the human coder by directly analyzing the pitch and amplitude data. Using a combination of PCA and LDA, we found further evidence that moral, prudential, and pragmatic transgressions elicit different vocal responses from mothers.

These findings support the notion that mothers' emotional signals play an important role in infants' early acquisition of norms (Dahl et al., 2011; Kochanska, 1994). Past research suggest that infants are generally sensitive to emotional signals from others (Mumme et al., 1996; Sorce et al., 1985; Vaish & Striano, 2004; Walle & Campos, 2012). By attending to their mothers' vocal tone, infants can come to understand that moral and prudential transgressions tend to be seen as more negative than pragmatic transgressions. Another strong indicator of the relative emphasis on moral and prudential transgressions is the

Table 4  
PCA-LDA Cross-Validation Results: Situation Type

True situation type	Predicted situation type		
	Moral	Prudential	Pragmatic
Moral	.53	.22	.25
Prudential	.28	.47	.25
Pragmatic	.21	.27	.52

Note. Numbers show respective mean proportions of vocalizations in the test sets that were classified as moral, prudential, and pragmatic by the LDA models. Cross-validation was done by randomly selecting 95% of vocalizations as a training set and using the remaining 5% as a test set, and repeating this procedure 100 times. PCA = principal component analysis; LDA = linear discriminant analysis.



*Figure 5.* Results of LDA for situation type. The upper right plot shows density contour plots for the three situation types plotted against the two linear discriminants. Contour lines are drawn at estimated densities starting at 0.12 (with 0.0075 increments). Dotted lines represent decision boundaries for the LDA classifier. Solid points show the center (median) points for each emotion category on the two linear discriminants. The marginal plots show the results of multiplying the matrix of eigenvectors for the 25 pitch PCs and the 9 intensity PCs with the vectors of corresponding scaling values for the first (bottom figures) and second (left figures) linear discriminants (calculated separately for pitch and intensity).

much higher response rates to these transgressions compared with response rates when witnessing pragmatic transgression (Study 2). Beyond mere valence and intensity, the mothers' emotional responses also differentiated between moral and prudential transgressions. In particular, the mothers' angry responses to moral transgressions can indicate to infants that mothers expect them to know better than to violate the very important norm against harming others, whereas the fearful responses to prudential transgressions signal to infants that they themselves are at immediate risk for harm.

The two studies presented here add to past research in several ways. First, past research on mothers' vocal communication and transgression-related interactions with infants has tended to see prohibitive messages as relatively homogenous (Fernald, 1989; Kochanska, 2002; Kochanska & Aksan, 2006; Trainor et al., 2000). The present studies, along with other research with older children and adults, demonstrate the need for distinguishing between types of transgressions (Dahl & Campos, 2013; Nucci & Weber, 1995; Smetana, 1989, 2013). Across the life span, people's understanding of the nature of a transgression is a

major determinant for how they think and act in the context of that transgression. A second difference between the present studies, and much of past research on infant-directed emotional vocalizations, is that we relied on naturally occurring and videotaped transgressions rather than verbal instructions for eliciting mothers' vocal responses. That is, the distinction between different types of utterances was never communicated to mothers, strengthening our conviction that moral, prudential, and pragmatic transgressions represent categories that, intentionally or unintentionally, guide mothers' spontaneous interactions with their children.

Finally, this research illustrates the benefits of combining naturalistic and experimental methods investigating the same question (Bandura & Walters, 1963; Gibson, 1979). The naturalistic data from Study 1 provides evidence that mothers use different emotional vocal responses to moral, prudential, and pragmatic transgressions in everyday life. Study 1 also allowed us to estimate the frequency of transgression-related vocalizations early in the second year of life. In contrast, Study 2 allowed us to improve the recording quality, eliminate confounds (such as other interventions co-occurring with the mothers' vocalizations), and investigate mothers' tendency *not* to respond to the three transgression types. The improved recording quality had benefits both for the human coding (allowing us to code the PL codes with good reliability) and for the quantitative analysis of the sounds signals. Here, we used a novel combination of PCA and LDA, allowing us to predict what kind of transgression type had elicited a given clip without reliance on either predefined vocal categories or an excessive number of correlated predictors. Beyond the scope of this article, these results provide additional evidence that the PCA-LDA approach used here is a promising way of analyzing human vocalizations (see Theunissen et al., 2013).

The present work calls for several extensions. It will be important to investigate how infants make use of differentiated vocal responses to transgressions. Past research gives reason to believe that *prototypical* facial and vocal emotional signals affect infant approach or avoidance of objects or locations (Miyake, Campos, Kagan, & Bradshaw, 1986; Sorce et al., 1985). However, the prototypic expression of an emotion, for instance, through the voice, does not always correspond to the expression of that emotion in a given situation (what Campos, Dahl, and He, 2010, refer to as the equifinality of emotional expressions). For this reason, infant use of the specific vocal tones documented herein must be studied directly. Moreover, children's acquisition of moral, prudential, and pragmatic norms involves more than avoidance of prohibited activities. Through the gradual construction of norms, children become able to evaluate the actions of others and to coordinate multiple concerns in multifaceted situations (Dahl et al., 2011; Hoffman, 2001; Turiel, *in press*). Additional research is needed to understand how mothers' spontaneous emotional reactions to transgressions affect infants' construction, adherence to, and application of moral, prudential, and pragmatic norms.

A second important extension of the present work is to study a greater variety of cultural groups. As stated in the introduction of this article, we did not expect significant effects of mothers' ethnic background on their vocal emotional responses because we included relatively small and homogeneous samples. Our

results were consistent with this expectation. This null finding, of course, should not be taken as evidence that there is no cultural variability in mothers' differentiated vocal responses to infant transgressions.

It is useful to distinguish between two potential sources of cultural variability: variability in parental beliefs and variability in expressive styles. For instance, certain communities do not consider it appropriate to tell young infants that they are doing something wrong, even if an infant is hitting someone (Briggs, 1974; Rogoff, 2003). In order to "toughen" their child, mothers of low socioeconomic status in Baltimore, observed by Miller and Sperry (1987), would encourage toddlers to hit back when someone had harmed them, whereas *unprovoked* aggression was discouraged. Such differences in beliefs about effective or appropriate parenting strategies appear to explain at least some cultural differences in parental behaviors toward infants (Harwood, Schoelmerich, Schulze, & Gonzalez, 1999).

Another, albeit related, source of variability is variability in expressive styles. Although there is evidence of some cross-cultural similarities in the expression and recognition of vocal emotion (Laukka et al., 2013; Papoušek, Papoušek, & Symmes, 1991; Sauter, Eisner, Ekman, & Scott, 2010), there are also cultural differences in the emotional expressions that are used or accepted in specific situations (Briggs, 1974; Elfenbein, Beaupré, Lévesque, & Hess, 2007; Howell, 1981). Some of these differences may be attributable to what Ekman and Friesen (1969) referred to as *cultural display rules*.

The present studies suggest that mothers' differentiated emotional vocalizations are a central feature of young children's everyday experiences with transgressions. Using multiple methods, we obtained consistent evidence that infant moral transgressions are associated with intense angry vocal tones, prudential transgressions are associated with fearful (and to some extent comforting) vocal tones, and pragmatic transgressions tend to elicit joyful and comforting vocal tones (or no response at all). As infants with limited linguistic abilities explore transgressive behaviors, these nuanced signals may prove crucial for children's early acquisition of norms.

## References

- Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. New York, NY: Holt, Rinehart, & Winston.
- Biringen, Z., Emde, R. N., Campos, J. J., & Appelbaum, M. I. (1995). Affective reorganization in the infant, the mother, and the dyad: The role of upright locomotion and its timing. *Child Development, 66*, 499–514. doi:10.2307/1131593
- Briggs, J. L. (1974). *Never in anger: Portrait of an Eskimo family*. Cambridge, MA: Harvard University Press.
- Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science, 296*, 1435. doi:10.1126/science.1069587
- Campos, J. J., Dahl, A., & He, M. (2010). Beyond breaches and battles: Clarifying important misconceptions about emotion. *Emotion Review, 2*, 100–104. doi:10.1177/1754073909355013
- Camras, L. A., Shuster, M. M., & Fraumeni, B. R. (2014). Emotion socialization in the family with an emphasis on culture. *Children and Emotion: New Insights into Developmental Affective Science, 26*, 67–80. doi:10.1159/000354355
- Cole, P. M., & Tan, P. Z. (2007). Emotion socialization from a cultural

- perspective. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 516–542). New York, NY: Guilford Press.
- Dahl, A., & Campos, J. J. (2013). Domain differences in early social interactions. *Child Development, 84*, 817–825. doi:10.1111/cdev.12002
- Dahl, A., Campos, J. J., & Witherington, D. C. (2011). Emotional action and communication in early moral development. *Emotion Review, 3*, 147–157. doi:10.1177/1754073910387948
- Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*, 49–65. doi:10.1111/j.2044-835X.1983.tb00543.x
- Dunn, J. (1988). *The beginnings of social understanding*. Oxford, UK: Blackwell. doi:10.4159/harvard.9780674330610
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49–98.
- Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion, 7*, 131–146. doi:10.1037/1528-3542.7.1.131
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*, 1–173; discussion 174–185.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497–1510. doi:10.2307/1130938
- Gärbling, A., & Gärling, T. (1995). Mothers' anticipation and prevention of unintentional injury to young children in the home. *Journal of Pediatric Psychology, 20*, 23–36. doi:10.1093/jpepsy/20.1.23
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Gralinski, J. H., & Kopp, C. B. (1993). Everyday rules for behavior: Mothers' requests to young children. *Developmental Psychology, 29*, 573–584. doi:10.1037/0012-1649.29.3.573
- Harwood, R. L., Schoelmerich, A., Schulze, P. A., & Gonzalez, Z. (1999). Cultural differences in maternal beliefs and behaviors: A study of middle-class Anglo and Puerto Rican mother-infant pairs in four everyday situations. *Child Development, 70*, 1005–1016. doi:10.1111/1467-8624.00073
- Hastings, P. D., & Grusec, J. E. (1998). Parenting goals as organizers of responses to parent-child disagreement. *Developmental Psychology, 34*, 465–479. doi:10.1037/0012-1649.34.3.465
- Hay, D. F. (2005). The beginnings of aggression in infancy. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 107–132). New York, NY: Guilford Press.
- Hoffman, M. L. (2001). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press. doi:10.1037/10433-003
- Honig, A. S., & Chung, M. (1989). Child-rearing practices of urban poor mothers of infants and three-year-olds in five cultures. *Early Child Development and Care, 50*, 75–97. doi:10.1080/0300443890500107
- Howell, S. (1981). Rules not words. In P. Heelas & A. Lock (Eds.), *Indigenous psychologies: The anthropology of the self* (pp. 133–143). London, UK: Academic Press.
- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.
- Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65–135). New York, NY: Oxford University Press.
- Kim, L. (2013). *Young children's understanding of pragmatic norms* (Unpublished honor's thesis). University of California, Berkeley, CA.
- Kochanska, G. (1994). Beyond cognition: Expanding the search for the early roots of internalization and conscience. *Developmental Psychology, 30*, 20–22. doi:10.1037/0012-1649.30.1.20
- Kochanska, G. (2002). Committed compliance, moral self, and internalization: A mediational model. *Developmental Psychology, 38*, 339–351. doi:10.1037/0012-1649.38.3.339
- Kochanska, G., & Aksan, N. (2006). Children's conscience and self-regulation. *Journal of Personality, 74*, 1587–1617. doi:10.1111/j.1467-6494.2006.00421.x
- Kuczynski, L., Kochanska, G., Radke-Yarrow, M., & Ginius-Brown, O. (1987). A developmental interpretation of young children's noncompliance. *Developmental Psychology, 23*, 799–806. doi:10.1037/0012-1649.23.6.799
- Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Chui, W., . . . Thingujam, N. S. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in Psychology, 4*, 353. doi:10.3389/fpsyg.2013.00353
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.
- Miller, P., & Sperry, L. L. (1987). The socialization of anger and aggression. *Merrill-Palmer Quarterly, 33*, 1–31.
- Miyake, K., Campos, J. J., Kagan, J., & Bradshaw, D. L. (1986). Issues in socioemotional development in Japan. In H. Azuma, K. Hakuta, & H. Stevenson (Eds.), *Kodomo: Child development and education in Japan* (pp. 239–261). San Francisco, CA: Freeman.
- Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development, 67*, 3219–3237. doi:10.2307/1131775
- Nucci, L., & Weber, E. K. (1995). Social interactions in the home and the development of young children's conceptions of the personal. *Child Development, 66*, 1438–1452. doi:10.2307/1131656
- Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meaning of melodies in motherese in tone and stress languages. *Infant Behavior & Development, 14*, 415–440. doi:10.1016/0163-6383(91)90031-M
- Power, T. G., & Parke, R. D. (1986). Patterns of early socialization: Mother-and father-infant interaction in the home. *International Journal of Behavioral Development, 9*, 331–341. doi:10.1177/016502548600900305
- Rijt-Plooij, H. H., & Plooij, F. X. (1993). Distinct periods of mother-infant conflict in normal development: Sources of progress and germs of pathology. *Journal of Child Psychology and Psychiatry, 34*, 229–245.
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford, UK: Oxford University Press.
- Ross, H. S. (1996). Negotiating principles of entitlement in sibling property disputes. *Developmental Psychology, 32*, 90–101. doi:10.1037/0012-1649.32.1.90
- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *PNAS Proceedings of the National Academy of Science of the United States of America, 107*, 2408–2412. doi:10.1073/pnas.0908239106
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227–256. doi:10.1016/S0167-6393(02)00084-5
- Shrivastav, R., Sapienza, C. M., & Nandur, V. (2005). Application of psychometric theory to the measurement of voice quality using rating scales. *Journal of Speech, Language, and Hearing Research, 48*, 323–335. doi:10.1044/1092-4388(2005)022
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3*, 365–394. doi:10.1207/S15327078IN0303\_5
- Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology, 25*, 499–508. doi:10.1037/0012-1649.25.4.499

- Smetana, J. G. (2013). Moral development: The social domain theory view. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology, Vol. 1: Body and mind* (pp. 832–864). Oxford, UK: Oxford University Press.
- Smetana, J. G., Kochanska, G., & Chuang, S. (2000). Mothers' conceptions of everyday rules for young toddlers: A longitudinal investigation. *Merrill-Palmer Quarterly, 46*, 391–416. doi:10.2307/23093738
- Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21*, 195–200. doi:10.1037/0012-1649.21.1.195
- Theunissen, F. E., Mouterde, S., & Mathevon, N. (2013). A neuroethological analysis of the information in propagated communication calls. *Proceedings of Meetings on Acoustics, 19*, 1–6.
- Tisak, M. S. (1993). Preschool children's judgments of moral and personal events involving physical harm and property damage. *Merrill-Palmer Quarterly, 39*, 375–390. doi:10.2307/23087427
- Tisak, M. S., Nucci, L. P., & Jankowski, A. M. (1996). Preschool children's social interactions involving moral and prudential transgressions: An observational study. *Early Education & Development, 7*, 137–148. doi:10.1207/s15566935eed0702\_3
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*, 188–195. doi:10.1111/1467-9280.00240
- Turiel, E. (in press). Moral development. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science* (7th ed., Vol. 1). Hoboken, NJ: Wiley.
- Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science, 7*, 261–269. doi:10.1111/j.1467-7687.2004.00344.x
- Walle, E. A., & Campos, J. J. (2012). Interpersonal responding to discrete emotions: A functionalist approach to the development of affect specificity. *Emotion Review, 4*, 413–422. doi:10.1177/1754073912445812

## Appendix

### Sample Equations and Description of Variables

#### Study 1 Example: GLMM for Presence of Stern Vocalizations

$$present_{ij} \sim \text{Bernoulli}(p_{ij})$$

$$\text{logistic}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \text{intercept}_j + \beta_{12} * \text{situation}_{ij}$$

$$\text{intercept}_j \sim N(\text{intercept}_0, \sigma_{\text{intercept}}^2)$$

$Present_{ij}$  indicates whether or not a stern tone was used by mother  $j$  in situation number  $i$ ;  $p_{ij}$  represents the probability of mother  $j$  using a stern tone in situation  $i$ ;  $\text{intercept}_j$  is the intercept for mother  $j$ ; and  $\text{situation}_{ij}$  indicates the situation type (moral, prudential, or pragmatic) for mother  $j$  in situation  $i$ .  $\beta_{12}$  is shorthand for the two situation coefficients corresponding to the two dummy-coded situation variables. The intercepts are modeled as normally distributed around  $\text{intercept}_0$  with a variance of  $\sigma_{\text{intercept}}^2$ .

#### Study 2 Example: GLMM for Presence of Stern Vocalizations

$$present_{ij} \sim \text{Bernoulli}(p_{ij})$$

$$\text{logistic}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \text{intercept}_j + \beta_{12} * \text{situation}_{ij}$$

$$\text{intercept}_j \sim N(\text{intercept}_0, \sigma_{\text{intercept}}^2)$$

$Present_{ij}$  indicates whether or not a stern tone was used by mother  $j$  in response to video clip number  $i$ ;  $p_{ij}$  is the probability of mother  $j$  using a stern tone for clip  $i$ ;  $\text{intercept}_j$  is the intercept for mother  $j$ ; and  $\text{situation}_{ij}$  indicates the situation type for mother  $j$  in clip  $i$ . The intercepts are modeled as normally distributed around a mean of  $\text{intercept}_0$  with a variance of  $\sigma_{\text{intercept}}^2$ .

Received October 8, 2013

Revision received February 14, 2014

Accepted February 25, 2014 ■