Development of perceptual expertise in emotion recognition

Seth D. Pollak, Michael Messner, Doris J. Kistler, Jeffrey F. Cohn

1. Introduction

Humans rely heavily on facial expressions when rapidly communicating emotional states to others. Although considerable attention has been devoted to how the nervous system responds to faces as compared with other objects, there are still many questions about factors that affect the perception of emotion-specific information communicated by the face. This experiment addresses how the perception of affective signals is influenced by children's early perceptual and social experiences. The findings suggest that perceptual learning influences emotion perception, and that social experience has an important role in calibrating properties of emotion perception according to the frequency or salience of affective expressions populating an individual's environment.

Physical differences between facial expressions are small compared with the range of variation in many cases of non-face objects. Yet young infants are sensitive to the subtle differences distinguishing facial expressions of emotion (Bornstein & Arterberry, 2003; Montague & Walker-Andrews, 2001). Debate continues about whether the human brain has a specialized neural subsystem for face perception (Farah, Wilson, Drain, & Tanaka, 1998; Kanwisher, 2000; Moscovitch, Winocur, & Behrmann, 1997) or whether these neural systems are developed through extensive expertise with faces (Diamond & Carey, 1986; Gauthier & Tarr, 1997). In either case, it is clear that over the course of development, perceptual learning involves interactions between areas involved in face recognition and those involved in spatial attention, feature binding and memory recall (Dolan et al., 1997). In addition, experience from other modalities, such as auditory information, likely contributes to children's understanding of facial expressions.

Several studies have demonstrated pronounced effects of visual experience on subsequent perceptual performance. The range of approaches used to address this issue includes low-level retinotopically specific effects (McKee & Westheimer, 1978) and single-unit recordings of inferior
temporal cortex (Fink et al., 1996; Heinze et al., 1994) to higher-order changes in object encoding (Carey & Diamond, 1994) and effects of culture (Elfenbein & Ambady, 2003). These data suggest that changes in categorization and automatization of perceptual processing occur based upon practice with specific classes of objects (Gauthier, Skudlarski, Gore, & Anderson, 2000; Tanaka, Curran, & Sheinberg, 2005). Such perceptual learning effects may also hold true for facial expressions of emotion.

Extant research suggests that the neural mechanisms underlying face perception can be altered by previous experience with faces (Leopold, O’Toole, Vetter, & Blanz, 2001; Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003; Webster, Kapling, Mizokami, & Duhameil, 2004; Webster & Maclin, 1999; Zhao & Chubb, 2001). Less is known about perception of discrete emotions. We are all exposed to facial expressions of emotion that are, in many ways, quite similar, making it difficult to evaluate the role of experience in the organization of emotion processing systems. Adults perceive basic emotional expressions as belonging to discrete categories (Young et al., 1997). In prior research we found that young children perceive emotional expressions based upon similar categories, but abused children—who have experienced frequent displays of extreme hostility—display broader perceptual categories of anger (but not other emotions) relative to nonmaltreated children (Pollak & Kistler, 2002). We also found that abused children could accurately identify angry faces based upon less perceptual information than typically developing children (Pollak & Sinha, 2002). These experiments suggest a role for perceptual learning in emotional development. However, most studies that examine the perception of emotional expressions in the face use static images that represent peak emotional expressions or morphed artificial images to approximate facial movement. Such stimuli carry important advantages including valid measurement of the emotion being conveyed. The emotional stimuli that children experience during interpersonal interactions are dynamic rather than static. Therefore, the present experiment was designed to test whether a feature of perceptual learning of emotions involves the ability to formulate valid predictions about changes in facial musculature during early stages of the formation emotional expressions.

Generating emotional expressions requires sequenced movements of facial musculatures (Ekman & Friesen, 1978). One hypothesis is that to correctly identify an emotion, observers use early information from these muscle movements to generate hypotheses about what emotion is being displayed. Once those characteristic movements have been extracted and mapped onto emotion categories, they provide additional social information. Therefore, rapid recognition of emotional signals facilitates adaptive social functioning. Observers integrate form and motion into judgments about facial information. This process appears to be guided by the observers’ expectations, perceptual sensitivity, and mental representations of emotional knowledge and previous experience/exposure (Knapmeyer, Thornton, & Bulthoff, 2003; see also O’Toole, Roark, & Abdi, 2002 for discussion).

In this study, we examine whether perceptual learning can shape the dynamics of emotion perception. To do so, we evaluated whether variations in social experience affect the perceptual judgments that observers made while viewing the realistic unfolding of human emotional expressions. It is not possible to experimentally manipulate an individual’s prior knowledge of basic facial expressions. Therefore, we tested children whose experiences with the communication of emotion may deviate in important ways from societal norms: children who had been severely physically maltreated by their parents. This is a population known to experience perturbations in both the frequency and content of their emotional interactions with caregivers as compared with children from nonmaltreating families (Pollak, 2008; U.S. Department of Health and Human Services, 2004). Reasoning that physically abusive episodes are likely marked by salient displays of anger, we predicted that abused children would begin to accurately recognize angry facial expressions earlier in the dynamic unfolding of the expression, reflecting a form of perceptual expertise in emotion recognition. However, because there is no a priori reason to believe that the information processing of these children generally differs from other children, we did not predict group differences in recognition processes for other emotional expressions. In sum, we expected that the effects of these children’s experiences would be reflected in a specific perceptual adaptation to the dynamics of expressions of anger.

2. Method

2.1. Participants

Ninety-five 9-year-old children participated in this experiment. Of these, 49 were physically abused (mean = 9.48 years, sd = 1.5 months; 45% non-Caucasian, 49% female). The comparison group included 46 children (mean = 9.50 years, sd = 1.6 months; 48% non-Caucasian, 50% female). An additional 80 adults and 80 9-year-old children participated in a preliminary study (described below) to rate the facial stimuli. None of the 160 individuals who participated in the rating studies were included in the perceptual experiment. Physically abused and non-abused comparison children were screened first through the registry of state Child Protective Services records for abuse histories and then in our laboratory with the Parent–Child Conflict Tactics Scale (Straus, Hamby, Finkelhor, Moore, & Runyan, 1998), a measure of the extent to which a parent has carried out specific acts of physical aggression toward the child. Non-abused comparison families were recruited from the community. Abused and control children had similar psycho-demographic characteristics: parents from both groups were unskilled/semi-skilled workers or were skilled manual workers. All children had normal or corrected-to-normal vision. To further support the claim that abused children develop within environments marked by higher levels of anger, parents completed the State-Trait Anger Expression Inventory (Spielberger, 1988), a self-report measure of subjective experiences and expression of anger.
2.2. Stimuli

Image data were from the Cohn–Kanade Facial Expression Database (Kanade, Cohn, & Tian, 2000), in which posers were video-recorded while performing directed facial action tasks (Ekman, 2007). These included both single facial actions (e.g., lowering the brows) and facial actions associated with displays of specific emotions. Image sequences were filmed using a frontal camera and digitized from SVHS video into 640 × 480 pixel arrays with 24-bit color resolution. Each sequence begins with a neutral expression and ends with the target actions. For the present study, we selected four sequences for each of five emotions. No poser appeared in more than one sequence. The sequences varied slightly in duration, with a mean of 19.85 frames (0.66 s) and standard deviation of 7.91 frames (0.26 s). To provide uniform duration, these video segments were down-sampled to 10 frames reflecting, respectively, 5, 11, 16, 20, 36, 59, 69, 75, 83, and 90% of the percentage of the elapsed sequence from absent-to-peak emotional expression. The final data set consists of twenty different models, selected to represent both genders and several ethnicities, posing five emotions: angry, happy, fear, sad, and surprise. Twenty-five percent of the models posing each emotion were non-Caucasian and children’s performance was not influenced by the relationship between the race of the model in the stimulus and the race of the child, all ps > 2.

Images from the video sequences were validated using multiple methods. First, two trained FACS coders independently scored each one. Second, two groups of adults completed a recognition task in which they labeled each sequence using either forced-choice or free response formats. Third, two groups of typically developing 7–12 year-old children did the same, labeling each sequence using either forced-choice or free response formats. A forced-choice labeling format was used so that the data could be directly compared to the ratings of Ekman and Friesen (1977) and a free response rating was used to provide more stringent recognition criteria (Widen & Russell, 2003). These data are summarized in Table 1.

2.3. Procedure

Each child viewed five emotions, each posed by four different models. Each emotion was presented as a series of ten images beginning with a neutral emotional expression through the peak expression (Fig. 1). Thus, each child responded to 200 total trials (5 emotions × 4 posers per emotion × 10 images per poser). Each image was presented for three seconds. Children were instructed to determine what emotion each model was feeling after hearing what was whispered into his/her ear. Children made forced-choice responses on a touch screen monitor after each image, selecting among: nothing yet, happy, sad, angry, afraid, or surprised. Presentation of each emotion was randomized across participants.

3. Results

Our analyses test how early in the unfolding of emotional expressions children accurately recognize an emotional expression. The time course of identification for each emotion was assessed by fitting group identification functions relating the probability of identification to the percentage of the elapsed emotion sequence (neutral through peak). The data were collapsed across models posing the expressions since there were no group differences for the models. Functions were fit to a Weibull distribution in MATLAB using a publicly available algorithm (Wichmann & Hill, 2001) following the approach used in Pollak and Kistler (2002). These functions are displayed in Fig. 2; the identification thresholds, defined as the image value corresponding to an identification probability of .5, and the 95% confidence intervals are indicated on each function. A repeated measures analysis of variance revealed a main effect of Emotion, F(4, 90) = 66.98, p < 0.001, η² = 0.49. This effect was further qualified by an interaction of emotion × maltreatment group × image F(36, 3348) = 0.05, η² = 0.24. Anger tended to be identified earlier by the abused children (control M = 5.3, SE = 0.22; abused M = 4.7, SE = 0.21, p = 0.06, η² = 0.037). The groups did not differ in recognizing other emotions, identifying happy (control M = 7.0, SE = 0.16; abused M = 7.0, SE = 0.15, p = 0.94) and sad (control M = 5.7, SE = 0.25; abused M = 5.7, SE = 0.24, p = 0.96) expressions earlier than fearful (control M = 3.8, SE = 0.21; abused M = 3.9, SE = 0.20, p = 0.74) and surprised (control M = 4.6, SE = 0.20; abused M = 4.8, SE = 0.19, p = 0.60) expressions. Analyses of children’s response errors indicated that neither group showed a response bias for selecting any particular emotion, F(4, 372) = 1.14, n.s.

An ancillary issue concerns the difficulty in measuring children’s emotional environments. To address this issue, we asked parents directly about the frequency and severity of the ways in which they discipline their children using the Conflict Tactics Scale. As expected, all of the non-abusive families scored less than 10 on this measure, reflecting verbal approbations and occasional and mild corporal punishment. In contrast, abusive families scored between 20–77, reflecting frequent and severe actions that include

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.91</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.98</td>
<td>0.95</td>
<td>0.88</td>
<td>0.65</td>
</tr>
<tr>
<td>Fear</td>
<td>0.88</td>
<td>0.65</td>
<td>0.67</td>
<td>0.52</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.92</td>
<td>0.86</td>
<td>0.91</td>
<td>0.82</td>
</tr>
<tr>
<td>Anger</td>
<td>0.93</td>
<td>0.69</td>
<td>0.81</td>
<td>0.81</td>
</tr>
</tbody>
</table>
harsh physical discipline, choking, scalding, and realistic threats to injure or abandon the child. Groups differed (control $M = 2.5$, $SD = 2.7$; abused $M = 40.0$, $SD = 16.6$) in this regard, $F(1,94) = 229.1$, $p < 0.001$. While parents are likely to under-report abusive behavior on these measures, it is unlikely that parents would systematically over-report these kinds of behaviors. Finally, because our hypotheses involved children’s experience perceiving anger, we asked parents to complete the State-Trait Anger Expression Inventory. Parents did not differ in their reports of state anger—that is, a measure of how they were feeling at the moment (control $M = 45.4$, $SD = 3.5$; abused $M = 46.3$, $SD = 5.2$). However, abusive parents reported more trait anger—that is, how generally they feel and overtly express anger (control $M = 43.9$, $SD = 4.8$; abused $M = 55.1$, $SD = 13.5$), $F(1,92) = 28.01$, $p < 0.001$. Parent trait anger was associated with more abusive acts towards children, $r = 0.35$, $p = 0.001$, $d = 0.71$, and children whose parents reported more abusive behaviors demonstrated more perceptual sensitivity to angry faces, $r = 0.23$, $p = 0.03$, $d = 0.48$.

4. Discussion

The present experiment examined children’s construction of emotion representations from fragmentary or partial information as models generated naturalistic facial expressions. We found that children who have been exposed to unusually high levels of anger were able to accurately recognize anger early in the formation of the facial expression of anger, when fewer expressive cues (such as activation of facial musculature) were available. However, these children performed similarly to controls when viewing other emotional expressions, suggesting that the abused children neither had a bias to select anger nor that the abused group generally performed or understood the task differently than did controls. These data address both basic issues about the development of emotion perception and also applied issues about the mechanisms through which child abuse is associated with behavioral problems in children.

4.1. Perceptual learning and emotion

It is difficult to precisely measure the role of social experience on the organization of developing perceptual systems ([Pollak, in press]. Within seconds of post-natal life, the human infant experiences a wealth of emotional input: sounds such as coos, laughter, tears of joy; sensations of touch, hugs, restraints, soothing and jarring movements; blurred images of faces with smiles, wide eyes; and a host of smells including pheromones. After a few days, let alone after months or years, it becomes a significant challenge to quantify an individual’s emotional experiences. Almost instantly, the maturation of the brain has been confounded by experiences with emotions. To address this problem, we tested children who have experienced social environments that are marked by high levels of anger and hostility. The features that make their environments atypical serve as approximations of how environmental variations in experience may affect development. The present data suggest that variations in experience with emotional facial expressions can induce different perceptual processing strategies and influence how and when emotions are perceived. Our stimuli were designed to capture naturalistic unfolding of emotional expressions, which resulted in anger expressions involving lip tightening rather than the open-mouthed expressions associated with shouting. It is
possible that abused children may be even better at detecting expressions of anger involving overt, open mouth displays. Future research might explore whether abused children may be particularly adept at detecting suppressed anger expressions as well.

It has been proposed that experience-independent neural structures facilitate attention to face-like objects, and such stimulus preferences lead to increased learning based on experience or create cortical specialization for face processing (For discussion, see Nelson, 2001; Pascalis et al., 2005; Tarr & Gauthier, 2000). Similar processes might also underlie emotion perception. Our results are consistent with theories that perception is a conjoint function of current sensory input interacting with memory and attentional processes (Kosslyn, 1996; Treisman, 1988), specifically, that memory and attention organize visual input through selective effects on emotion processing. Brain mechanisms appear to learn to perceive and discriminate stimuli on the basis of acquired behavioral significance. This suggests that perceptual learning involves both category-specific extrastriate neuronal populations and interactions with higher-order brain regions including circuitry of the amygdala, fusiform, lateral parietal cortex, and higher-order visual regions that coordinate this sort of high level perception (Dolan et al., 1997; Pegna, Khateb, Lazeyras, & Seghier, 2005).

4.2. Perceptual learning and psychopathology

Adaptation to their environments, in the form of perceptual learning, likely facilitates children’s efficient coding of important stimulus features from the social environment. How might environmental deviation—such as extremely hostile affective experiences—calibrate the processes related to emotion perception? The development of increased perceptual sensitivity for the fine-grained details of variation in affective expressions may provide a behavioral advantage for children living in threatening contexts, allowing earlier identification of salient emotions. Anger-related cues may become especially salient to physically abused children because they are associated with imminent harm. Early identification of an emotion on the basis of minimal visual information, if accurate, would be adaptive for physically abused children, therefore they may learn to make decisions about the signaling of anger using minimal visual information. This does not suggest that there is anything inherently problematic or unhealthy about anger; on the contrary, the argument is simply that children learn to use highly predictive signals through increased attention to minimal signals. In this manner, maltreatment may sensitize children to certain emotional information that may be adaptive in abusive contexts but maladaptive in more normative interpersonal situations (Pollak, 2008).

The present data highlight the important function of facial expressions in the social communication of changes in affective states. Much less is understood about how humans learn to process the temporal affective cues integral to real-life social exchanges. It appears that children who have experienced direct physical harm become sensitive to the social signals of anger. The implication of this finding is that the social relevance of facial movements affects processing and recognition of emotional expressions.
References


