Personal familiarity influences the processing of upright and inverted faces in infants

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.3389/neuro.09.001.2010</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:13548126">http://nrs.harvard.edu/urn-3:HUL.InstRepos:13548126</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
INTRODUCTION

Selectivity is a key feature of mature representations of facial appearance. As demonstrated by a substantial body of behavioral results obtained from adult observers, faces are processed and recognized by distinct processes that do not appear to be readily applied to control stimuli (including inverted, scrambled, or contrast-negated faces). Such features of face processing include high-fidelity encoding of so-called “2nd-order” features of the face (Rhodes et al., 1993) and holistic processing of face patterns as revealed by phenomena like the composite-face effect (Young et al., 1987) and the part-whole effect (Tanaka and Farah, 1993). That these features of adult face processing are generally evident in unaltered faces, but absent or reduced in faces that have been transformed by disorientation or other image-level manipulations indicates that the visual system has learned important distinctions between classes of face-like stimuli. That is, true faces are distinguished from non-face images that share many of the same low-level properties. Presently, we examine how such distinctions are acquired (and thus how selective representations are constructed) by investigating the selectivity of infants’ neural response to faces belonging to distinct categories.

The perceptual processes that contribute to face recognition in infancy become increasingly selective during the first year of life. Specifically, infants show poorer recognition of face exemplars within categories they are not exposed to frequently. This process has been termed “perceptual narrowing” (Nelson, 2001; Scott et al., 2007) and has been observed in the context of face categories defined by species (Pascalis et al., 2002; Sugita, 2008), gender (Quinn et al., 2002), and race (Kelly et al., 2005, 2007). In each case, younger infants are able to successfully individuate faces in a broader set of categories than older infants, with older infants being worse at discriminating between faces belonging to infrequently encountered categories. Perceptual narrowing results demonstrate that experience shapes selectivity during infancy, leading to profound perceptual consequences. The first year of life (specifically the 6- to 9-month age range) is thus a critically important time period for studying how distinctions are made between faces and non-face control stimuli, as well as between distinct categories of faces defined at varying levels of granularity.

Our aim in the present study was to characterize how distinct types of face selectivity may interact during this period of development. This is distinct from any effort to characterize true perceptual narrowing, but shares the same broad goal of perceptual narrowing research insofar as we wish to more fully understand the developmental timecourse of differential processing for various kinds of faces and face-like stimuli. Specifically, we ask how differential processing of faces at a sub-ordinate level (personally familiar vs. unfamiliar face) may lead to differential processing at a basic level (face vs. inverted face).

What do we mean when we say that faces may be distinguished at different levels of granularity or “scales”? Intuitively, we suggest that degrees of face selectivity can be considered hierarchically. A coarse face processing strategy (that is not very selective) may only differentiate between faces and non-faces (such as inverted, photo-negative or distorted faces (Figure 1, top row). A more sophisticated and selective representation may distinguish between face categories on the basis of species, gender, race, or age (Figure 1, middle row). Finally, a very sophisticated representation of facial appearance may be able to make distinctions between individual exemplars within a category. This naïve ontology is
Face recognition performance suffers more than generic object recognition following a 180-degree planar rotation of the image (Yin, 1969), and ERP waveforms also exhibit a marked inversion effect. Specifically, the N170 component in adults, for example, is initially driven by individual exemplars, while the latter would suggest that selectivity has been acquired and applied generally to new faces (Farroni et al., 2005), an ability that improves dramatically over subsequent months (Turati et al., 2004). Further, within the first few months infants can visually discriminate the mother’s face from that of a stranger (Pascalis et al., 1995; de Haan et al., 2001). Given that infants of sufficient age can differentiate between faces along these dimensions, how do familiarity and orientation interact at the level of the neural response? We used event-related potentials (ERPs) to examine the extent to which face category membership defined at one “level” (as in Figure 1) modulated the selectivity of the neural response at another level.

We varied face orientation (upright vs. inverted) in a within-subjects design, varying face familiarity across participants. Our index of the neural response to faces was the P400 component, which is a face-sensitive response observed over occipital electrode sites (de Haan and Nelson, 1999; Halit et al., 2003). This component begins to emerge as early as 3 months of age (Halit et al., 2003, 2004; Scott and Nelson, 2004; Macchi Cassia et al., 2006) and is thus a useful way to examine neural face processing during the first year of life. Our question is whether or not the P400 exhibits differential processing for upright and inverted faces as a function of face familiarity, or if there appears to be no influence of face familiarity on the magnitude of the inversion effect. The former case would suggest that face orientation selectivity for this component is initially driven by individual exemplars, while the latter would suggest that selectivity has been acquired and applied generally to wider subsets of faces by 6 months of age. We find evidence of an intriguing interaction between these levels of face categorization, suggesting that the exquisite selectivity exhibited by adult observers may be built up in an exemplar-driven fashion.

**MATERIALS AND METHODS**

**PARTICIPANTS**

Infants were recruited from a metropolitan area. All infants were born full-term with no known pre- or perinatal complications. Informed consent was obtained from the parents and infants were assigned to one of two experimental groups. Infants belonging to the “Mother’s Face” group were presented with their mother’s face during the task, and infants belonging to the “Stranger’s Face” group were presented with a stranger’s face (see Stimuli for details). The final sample in the “Mother’s Face” group consisted of 20 participants (8 males) with a mean age of 183 days (range 177–191 days). An additional 40 infants were tested but not included in the final sample due to eye and/or body movements that resulted in excessive artifact (n = 29) or fussiness that resulted in too few trials being recorded (n = 11). The final sample in the “Stranger’s Face” group consisted of 14 participants (9 males) with a mean age of 184 days (range 177–188 days). An additional 13 infants were tested but not included in the final sample due to excessive eye and/or body movements that resulted in excessive artifact (n = 11) or fussiness.
that resulted in too few trials being recorded \((n = 2)\). We emphasize that participants were only excluded for reasons relating to data quality (ERP artifacts) and not due to boredom or disinterest in the task. Further, our attrition rate (though high) is consistent with the 50–75% dropout rate typically reported for infant ERP studies (DeBoer et al., 2007).

**STIMULI**
The stimuli consisted of color images of female faces displaying neutral expressions. Each woman was photographed wearing a gray scarf in order to conceal any clothing while seated in front of a gray background. Infants in the “Mother’s Face” group viewed pictures of their own mother’s face presented multiple times in both upright and inverted orientations. Infants in the “Stranger’s Face” group viewed pictures of an unfamiliar female face (mothers of “Mother’s Face” infants) that was judged to be dissimilar-looking to their own mother, but matched for the presence of eyeglasses and/or race of the model. This face was also presented in upright and inverted orientations.

**PROCEDURE**
Testing took place in a dimly lit, quiet room after application of the sensor net. Infants were tested while sitting on their parent’s lap. Stimuli were presented using E-Prime software v1.2 (Psychology Software Tools Inc., Pittsburgh, PA, USA). The faces were presented on the center of the screen on a white background. The computer monitor was 48 cm wide and 31 cm high. When viewed from a distance of 60 cm, the faces on average subtended a horizontal angle of approximately 16° and a vertical angle of 15°. A video camera mounted above the monitor and centered on the infant’s face allowed for observation of the infant at all times during the testing session. On-line judgments were made to present the pictures only when the infant was attending to the monitor. Trials were immediately marked for deletion by the experimenter if the infant looked away during stimulus presentation.

Stimuli were presented for 500 ms followed by an experimenter controlled inter-stimulus interval of at least 1500 ms during which time the screen was white. The two orientations were randomly presented with equal probability, with the constraint that stimuli from the same orientation were not repeated more than three times in succession. Stimulus presentation continued until the infant became bored or too fussy to attend, with a maximum of 100 trials. The average number of total trials viewed by infants was 71 and 80, for the “Mother’s Face” and “Stranger’s Face” groups respectively.

**ERP RECORDING AND ANALYSIS**
ERPs were recorded using a 64-channel Geodesic Sensor Net v2.0 (Electrical Geodesics Inc., Eugene, OR). EEG was recorded continuously and referenced to a single vertex electrode (Cz). Signals were amplified using an EGI NetAmps 200 amplifier with a band-pass filter of 0.1–80 Hz and a sampling rate of 200 Hz. Impedances were checked on-line prior to beginning the session and were considered acceptable if lower than 50 KOhm.

Continuous EEG data were processed offline using NetStation v4.1.2 (Eugene, OR). A 30-Hz lowpass filter was applied and trials were constructed that consisted of a 100 ms baseline period and 1500 ms period following stimulus onset. Data were baseline corrected to the average voltage during the 100 ms prior to stimulus onset. Segmented data were edited for EOG and motion artifact. Data from individual sensors were rejected if there was artifact resulting from poor contact or movement. The entire trial was excluded if more than nine sensors had been rejected, or if an eye-blink or other significant artifact had occurred. Of the remaining trials, individual channels containing artifact were replaced using spherical spline interpolation. Individual subject averages were constructed separately for the upright and inverted faces \((M = 22\) trials per condition in both groups), and data were re-referenced to the average reference.

Inspection of the grand-averaged waveforms revealed a well-defined P400 component that was subsequently analyzed within a time window 345–599 ms. Electrode groupings and time windows were selected based on previous reports of this component and through visual inspection of the grand-averaged and individual waveforms. Ten occipital electrodes were identified for our analysis of the P400 \((32, 33, 36, 37, 38, 39, 40, 41, 44, 45)\) which were further partitioned into left, midline and right regions for analysis (Figure 2A).

**RESULTS**
The peak amplitude (maximum value attained within the previously mentioned time window) and latency of the P400 was analyzed for each group separately using a 2 × 3 repeated measures analysis of variance (ANOVA) with Orientation (upright, inverted) and Region (left, mid, right electrode groupings) as within-subjects factors. Greenhouse-Geisser corrected degrees of freedom were used and post hoc paired t-tests were conducted using a Bonferroni correction for multiple comparisons.

"MOTHER’S FACE” GROUP

**Amplitude**
There was a main effect of Orientation, \(F(1, 18) = 5.83, p = 0.027,\) such that the mother inverted \((M = 15 \mu V, SD = 5.6)\) elicited a significantly larger P400 peak amplitude than mother upright \((M = 12.6 \mu V, SD = 6;\) see Figure 2B). We note that in particular that there were no interactions of Region (left, midline, or right sensor groups), which is why we have chosen to present an average waveform over all 10 electrodes in Figure 2.

**Latency**
For latency to the P400 peak, there were no main effects or interactions of Region or Orientation.

"STRANGER’S FACE” GROUP

**Amplitude**
For the amplitude of the P400, there were no main effects or interactions of Orientation or Region. (see Figure 2C).

**Latency**
For latency to the P400 peak, there was a main effect of Region, \(F(2,12) = 5.23, p = 0.019.\) Paired comparisons revealed that the midline electrode group \((M = 478 ms, SD = 39)\) peaked marginally faster than the left electrode group only \((M = 506 ms, SD = 34), t(13) = 2.67, p = 0.057\).
In summary, we find evidence of a differential response to upright and inverted faces only in the “Mother’s Face” group. Specifically, the amplitude of the P400 is larger to inverted faces than to upright faces. In Figure 3, we display the amplitude data across subjects as a function of stimulus type, sensor region, and subject group. There were no effects of face orientation on latency in either group, which may be due to fairly large variability across subjects (pooled SD across all conditions and groups = 36 ms).

Given the difference in the number of subjects included in the “Mother’s Face” and “Stranger’s Face” group, we also conducted an additional Monte Carlo analysis to rule out the possibility that a significant effect was observed in the larger group (“Mother’s Face”) simply due to the larger sample size. We carried out a “jackknife” procedure in which 14 participants were randomly sampled with replacement 10,000 times from the full set of 19 participants in the Mother’s Face group. At each iteration, the $F$-statistic for the main effect of face inversion was computed. This allows us to compare the distribution of $F$-ratios obtained from this procedure to the observed $F$-ratio obtained from the Stranger’s Face group. We find that this procedure yields an $F$-value for face inversion that is both significantly larger than one and also significantly larger than the $F$-ratio obtained from the Stranger’s Face group as determined by the 5th quantile of the empirical jackknife distribution. We therefore conclude that the observed difference between the groups is likely not due to the difference in sample size.

DISCUSSION

Infants in the current study were presented with digital images of either their mother’s face or a stranger’s face presented in an upright and inverted orientation. ERP data revealed that the amplitude of the P400 was greater to mother inverted vs. mother upright. However, no amplitude differences for this component were observed for the upright vs. inverted stranger manipulation. In terms of our initial question regarding how face selectivity is realized across different conceptual scales, even as early as 6 months, evidence for exemplar-driven influence on face processing can be observed. We take this as evidence that face selectivity may indeed be learned in an exemplar-driven fashion. That is, increasing selectivity for facial appearance emerges by repeated exposure to particular individuals, resulting in an enhanced capacity for differential processing at multiple scales of selectivity.

Our result is consistent with adult data from a number of studies (Bruce, 1986). Most directly, Balas et al. (2007) reported a familiarity effect for judging face orientation, suggesting that increased exposure to a face during adulthood facilitates orientation processing. Along other dimensions of face variation, including gender (Rossion, 2002), race (Bruyer et al., 2004), and expression (Kaufmann and Schweinberger, 2004) similar familiarity effects are obtained. Narrowing may thus be driven by and constrained by the individual faces that are seen most often. It is also noteworthy that increased training with non-face categories including...
“Greeble” stimuli (Gauthier and Tarr, 1997) and texture patterns (Husk et al., 2007) leads to inversion effects in adults. This may mean that selectivity in any domain may be driven primarily by the individual items experienced which subsequently cut deeply across scales of stimulus categories.

How do our data relate to other studies of the face inversion effect in infancy? A key point of comparison is to consider our results alongside those reported by de Haan et al. (2002), who presented 6-month-old infants and adults with upright and inverted images of human and monkey faces. While they report a species-specific inversion effect on the N170 component in adults, they also report the absence of any inversion effects on the “infant N170” (a negativity between 200–350 ms) and a non-specific inversion effect that is evident at the P400. This latter result is not consistent with the data we have reported here, insofar as monkey faces should certainly be unfamiliar to the infants in their task, and the models used in the human face trials were also not familiar to the infants. In neither case would our data suggest that an inversion effect should be evident at the P400. Moreover, there is an additional important discrepancy between our results and this previous study: The inversion effect reported at the P400 in this previous study goes in the opposite direction as our inversion effect. Inverted faces (both human and monkey) elicit a smaller positive peak at the P400, and effect which would appear broadly consistent with other results concerning the “narrowing” of face representations. Can we reconcile our results with this previous report? It is difficult to speak conclusively about the mutual relationship between such distinct results, but we briefly point out two important points that may be relevant to the differences we have described above. First, we note that while we did not observe any topographical differences between conditions, the de Haan paper reports that the inversion effect for Monkey faces was only evident over the left hemisphere, but obtained over both hemispheres for human faces. Second, it is perhaps important that the present study used only one model face per child (presented multiple times in each orientation) while the de Haan, Pascalis and Johnson study reports the use of 20 model faces per condition. The use of more model faces is laudable, and strengthens the generality of their results. In our case, obtaining a wide range of highly familiar faces was not practical, constraining the number of faces we could use in the “Stranger’s Face” condition. We suggest that it is possible that the repetition of the same model face repeatedly may have important consequences on the components we have examined here, an issue that has not yet been studied in great depth.

Finally, what are we to make of the differing directions of the face inversion effects reported here and in the de Haan study? If we are to begin by comparing the results of each study to the established pattern of inversion effects in the adult literature we must decide if the important feature of the adult face inversion effect at the N170 is the fact that it is of greater absolute amplitude, or the fact that it is more negative. The former would appear to be consistent with our result, while the latter would be more consistent with an inversion effect in the opposite direction (a more negative positive deflection implies a smaller peak). One way to potentially resolve this ambiguity is to consider data from other components used to study adult face processing. Both the P1 and the P2 exhibit larger amplitudes for inverted faces (Itier and Taylor, 2002) suggesting that perhaps the hallmark of inversion effects is an overall increase in the absolute amplitude of the ERP signal, rather than increased negativity. Though this provides a tentative suggestion that our data are broadly consistent with adult ERP results, future study of the consistency of inversion effects at the P400 during the first year of life will ultimately prove invaluable. Finally, we note that changes in the sign of the inversion effect have been observed in longitudinal studies of infant memory (Bauer et al., 2006), indicating that the same child may “flip” the inversion effect between test sessions. This suggests that it is possible to find inversion effects of differing sign within a relatively short timespan of development.

We close by suggesting potential extensions of this study that would further clarify the consequences of differential processing at multiple scales of face categorization. First, an important extension of this work (that is unfortunately difficult to achieve due to limitations on infant ERP study designs) is the inclusion of multiple stimulus categories representing a more comprehensive sampling of “face space”. (Valentine, 1991). Our comparison of familiar to unfamiliar female face processing across distinct subject groups is a fairly crude tool for examining the rich multi-dimensional code for facial appearance. As we have already noted, our desire to use personally familiar faces severely necessarily limited the number of distinct exemplars we could present to our infant subjects with, potentially limiting the generality of our results. Should exploring a larger set of stimuli within individual subjects prove practical (perhaps including father’s faces, or other siblings), further elaboration of interactions across “scales” of face specificity would be vital to understanding the development of face recognition. Specifically, further characterizing how categories defined by race, age, or gender all interact with personal familiarity or basic-level face/non-face distinctions would be a crucial step towards describing the acquisition of selective face responses. Second, a similar means of extending the present results would be to consider the impact of personal familiarity (or any other categorical distinction between faces) on the processing associated with other kinds of transformed face stimuli. Here, we have only examined the inversion effect as one example of a “coarse-scale” distinction between face and non-face control images. There are, of course, a wide range of similar transformations that can be carried out on face images that leave many low-level features of the face intact, including color or luminance negation (Galper, 1970; Kemp et al., 1990), or scrambling of face parts. Indeed, many other aspects of face recognition (such as the acquisition of view-invariant recognition) are developing during the first year of life, making them interesting candidates for a similar analysis. Examining how differential processing at the level of exemplars or other face categories impacts differential processing of these other kinds of transformation would represent an additional important contribution.

Finally, we also point out that in our current data we were unable to consistently identify the so-called “infant N170,” an additional face-sensitive component that is frequently reported at ~290 ms (de Haan et al., 2002). Should it prove possible to obtain a sufficiently robust response that this component could also be analyzed in the same manner as we have examined the P400 here, it would be useful to determine the extent to which these two components index similar or distinct aspects of face selectivity behave similarly. Similarly, in terms of further understanding the neural
substance supporting our effects, limits in the spatial resolution of ERPs preclude us from drawing inferences as to where in the brain these processes are taking place. Extending the study presented here using Near-Infrared Spectroscopy (NIRS) would permit more robust localization and provide an important complement to this work (Otsuka et al., 2007).

To conclude, our findings suggest that by 6 months of age, the neural architecture underlying face processing reflects selectivity learned from experience with particular faces. Differential processing at coarse scales appears to be modulated by differential processing at comparatively finer scales. Overall, these data complement the extensive behavioral literature on the development of face processing by revealing that the ability to process information about familiarity and orientation are not fully independent.

ACKNOWLEDGMENTS

The writing of this paper was supported by grants from the NIH (NS32976 and MH078829) to Charles A. Nelson. The authors thank Venessa Peña and Hannah Mandel for their assistance in data collection and coding; and the parents who participated.

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any com- mercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 July 2009; paper pending published: 08 September 2009; accepted: 07 January 2010; published online: 22 February 2010.


Frontiers in Human Neuroscience www.frontiersin.org February 2010 | Volume 4 | Article 1 | 6