REPORT

The face behind the mask: a developmental study

Nim Tottenham,¹ Andrew C. Leon² and B.J. Casey¹

¹. Sackler Institute for Developmental Psychobiology, Weill Medical College of Cornell University, New York, USA
². Departments of Psychiatry and Public Health, Weill Medical College of Cornell University, New York, USA

Abstract

Faces are a rich and available source of social information, and the representation for faces is robust in adults (i.e. the face detection effect; Purcell & Stewart, 1988). The current study compared the developmental trajectory of the robustness of face perception against the trajectory for a non-face object. Participants (5–35 years old) were presented with rapid (17 and 33 millisecond) presentations of face and house stimuli and were instructed to identify the object category of the stimulus (face or house). There was an interaction between object type and age such that the developmental slope for face identification was steeper than the slope for house identification for the 17-millisecond presentation. These data show that faces are processed in a different way than a non-face object during the period from middle childhood through adolescence and adulthood, and this differential processing may involve the massive amount of exposure we have to faces.

Introduction

Face perception is one of the most well-developed visual skills that human adults possess. What does it mean that this skill is well developed in adults? For one, people use expert-level strategies when viewing faces that they do not use for most object categories. Manipulations that disrupt the faceness of the face-image eliminate the expert-performance associated with faces (Farah, Wilson, Drain & Tanaka, 1998; Yin, 1969). Likewise, strategies that typically aid face perception interfere with processing manipulated faces (e.g. composite effect (Young, Hellywell & Hay, 1987) and the face superiority effect (Homa, Haver, & Schwartz, 1976)). Another dimension of expertise is the strength of the representation. Expert face processing is accompanied by robust perceptual representations (Haxby, Gobbini, Furey, Ishai, Schouten & Pietrini, 2001) that support rapid recognition of faces (Purcell & Stewart, 1988). Even when faces and non-face objects are matched on luminance, contrast, brightness and spatial frequency, faces are identified more quickly than non-face images (Reinders, den Boer & Buchel, 2005). The current study examines the developmental trajectory of rapid face processing by measuring identification of faces and objects at speeds that lie at the threshold of visual awareness.

Tracking the developmental trajectory of face processing is difficult for two reasons. Most studies do not sample across ages from childhood through adulthood, but instead focus on single age groups (e.g. infants, children or adults). Secondly, there are multiple ways to process a face, which obscures the path of the trajectory. For example, it is clear from most developmental studies on face processing that featural processing is a skill used early in life and precedes the use of configural strategies (Diamond & Carey, 1977; Taylor, Edmonds, McCarthy & Allison, 2001). The literature is mixed as to when configural or holistic processing strategies emerge. Depending on the paradigm used, these skills have been suggested to emerge in late adolescence (Itier & Taylor, 2004), by 10 years (Carey & Diamond, 1977; Diamond & Carey, 1977; Mondloch, Geldart, Maurer & Le Grand, 2003), 7 years (Gilchrist & McKone, 2003), 6 years (Carey & Diamond, 1994; Flin, 1985; Tanaka, Kay, Grinnell, Stansfield & Szechter, 1998), 5 years (Brace et al., 2001), 4 years (Freire & Lee, 2001; Pellicano & Rhodes, 2003), and even suggested to be available by 7 months (Cohen & Cashon, 2001), although the chances of children relying on this tactic are low relative to their reliance on featural processing (Schwarzer, 2002). Despite how methodologically constrained these studies are, the lack of overlap in findings makes it difficult to see the developmental trajectory of face processing.

Most behavioral studies examining the development of face processing throughout childhood have done so by testing discriminations for face identity, or the ability to
make face–face distinctions (with the exception of Carey & Diamond, 1977; Schwarzer, 2002). It is impossible using such methods to tell if the changes in face processing skills are specific to faces or whether they reflect general development of visual object processing. Examining face–object discriminations (or first-order discriminations (Maurer, Le Grand & Mondloch, 2002)), an approach used routinely in neurocognitive studies (e.g. Ishai, Ungerleider, Martin & Haxby, 2000; Ishai, Ungerleider, Martin, Schouten & Haxby, 1999; Kanwisher, McDermott & Chun, 1997; Puce, Allison, Asgari, Gore & McCarthy, 1996; Sergent, Ohta & MacDonald, 1992), offers more compelling evidence for processes that underlie perception of faces over other object categories.

Face processing appears to be different from non-face object processing early in life, but remains immature until later in development. At birth, face-like arrangements are generally preferred over other arrangements (Johnson, Dziurawiec, Ellis & Morton, 1991). This preference for face-like stimuli may have less to do with faces per se and more to do with a high spatial frequency pattern in the upper half of the stimulus (Simion, Cassia, Turati & Valenza, 2001; Turati, 2004; Turati, Simion, Milani & Umlita, 2002). Faces have been shown to be differentiated from other highly familiar objects electrophysiologically by 6 months (de Haan & Nelson, 1999). Older children (2–5 years), despite being able to differentiate between faces and other objects, use unsophisticated strategies for processing both object types (i.e. featural processing; Schwarzer, 2002). A more sophisticated visual strategy develops during late childhood, such that there is an increasing tendency to show a face inversion effect (laboratory evidence of visual expertise) from 6 to 10 years old, but that is not so for non-face objects such as houses (Carey & Diamond, 1977).

One means of delineating the developmental trajectory of face processing skills beyond examining strategy-based processing is measuring speeded object identification, which reveals more about the robustness of the perceptual representations of faces at the initial stages of processing (Reinders et al., 2005). Speeded face–object discriminations have only been studied in adult populations and non-human primates. Faces are more efficiently processed relative to non-face objects (e.g. Purcell & Stewart, 1988; Reinders et al., 2005; Shelley-Tremblay & Mack, 1999). Electrophysiological studies show that faces are distinguished from other objects early in the visual stream, within the first 200 msec (Eimer, 2000; Tanaka & Curran, 2001). Single-cell recordings in the macaque show that face-selective neurons in superior temporal sulcus and inferior temporal cortex respond to face stimuli presented as briefly as 20 ms, when a visual mask interrupts processing (Rolls, 2004; Rolls & Tovee, 1994; Rolls, Tovee & Panzeri, 1999). Such rapid processing suggests that the neural representation of faceness is robust and therefore, recognition is facilitated (Reinders et al., 2005).

Behaviorally, speed of processing differences between faces and objects have been captured by the ‘face-detection’ effect (Purcell & Stewart, 1988). Non-face objects, such as inverted faces, rearranged faces (Purcell & Stewart, 1988), letters and random dots (Hoshiyama, Kakigi, Watanabe, Miki & Takeshima, 2003), all require longer presentation times to be processed than do faces. This cannot be attributed to low-level stimulus features, such as luminance and spatial frequency, as the difference remains even when controlling for these potential confounds (Reinders et al., 2005; Shelley-Tremblay & Mack, 1999). Instead, because faces represent a highly meaningful source of information and consequently one with which humans have extended experience, faces are more readily processed by the visual system than non-face objects. The face detection effect is tested with a stimulus detection task (present vs. absent). The current study examines stimulus identification instead, with faster identification of faces being referred to as a ‘face advantage effect’.

This study uses backward masking, which is an appropriate task for all age groups, to examine face–object discriminations from childhood through young adulthood. Whether face skills have a different developmental slope than a non-face object is tested. Characterizing the developmental trajectory for the face advantage effect, which reflects the robustness of the face representation and marks a mature sensitivity to first-order relations of faces, is an overarching goal of this study. We hypothesized that faces would be more easily recognized than houses, and this advantage for processing faces would be enhanced with development.

Method

Participants

Ninety healthy, right-handed participants aged 5–12-years-old (n = 44, 24 females, mean age = 8.3, SD = 2.1), 13–17-years-old (n = 18, 10 females, mean age = 15.4, SD = 1.5) and 18–39 years old (n = 28, 14 females, mean age = 26.4, SD = 4.5) or their parents gave written informed consent. All subjects met our inclusion criteria of no history of neurological or psychiatric illness and an IQ above 80. Subjects were paid for their participation.

Apparatus and materials

Images were presented with E-prime® software on a 450 MHz PC computer with a 60-Hz monitor. Digital grayscale
images of faces came from the NimStim Face Set (Tottenham, Tanaka, Leon, McCarr, Nurse, Hare, Marcus, Borscheid, Casey & Nelson, submitted). Face images included male and female European-American and African-American models. Stimuli included expression and expressionless faces. Digital grayscale images of houses were Microsoft® Clipart images that were selected for similarity in size and luminance. Luminance was equated across all stimuli. Masks were visual noise masks that lasted 500 msec. All images were presented at a 10.75° vertical visual angle and an 8.21° horizontal visual angle.

Procedure

Face and house images were presented in random order, each with the mask following immediately (see Figure 1). Presentation rates were 17 and 33 msec. After each mask, subjects were instructed to indicate verbally whether they saw a house, a face, or whether they saw nothing (i.e. ‘miss’). They were asked to guess if they could. Responses were recorded by the experimenter. Each run had an equal number of face and house trials. There were 40 trials per run. Some of the subjects asked to stop early because of boredom or fatigue; thus subjects completed at least two runs and no more than four runs. Because each run had equal number of face and house trials and because subjects did not stop before completing a run, all subjects were presented with an equal number of face and house trials. Participants received no feedback following trials.

Data were aggregated to create percent identification scores, where false alarms and errors were subtracted to control for any response bias, thus resulting in an approximation of d-prime. D-prime accounts for response biases and is therefore preferred over simply counting correct responses (Itier & Taylor, 2004). However, because three response options were available (house, face or don’t know), it was impossible to precisely calculate d-prime. Therefore, we employed a different means for accounting for response biases. For example, a face identification score was calculated by subtracting errors on face stimuli (reported ‘house’ or miss) and false alarms for faces (a house was presented and subject reported face) from the total number of trials for face stimuli. This value was divided by the total number of face stimuli presented and then multiplied by 100. The same calculation was performed for house stimuli. While this type of calculation may over-account for the number of false alarms a subject could make, we would rather be conservative and rely on a calculation that accounted for this response bias since there may be biases at different ages for reporting one object class (e.g. faces). Four subjects (three adults, one 5-year old) were influential outliers (based on the standardized difference of beta) and were excluded from analyses. The scores over repeated exposures to face and house stimuli are highly correlated (intraclass correlation for 17-msec presentation = .68 and intraclass correlation for 33-msec presentation = .77; number of observations = 86); therefore mixed-effects linear regression models (Laird & Ware, 1982) were examined for percent identification (the dependent variable) separately for the 17-msec presentation and for the 33-msec presentation. Fixed effects in these models included object type (face or house) and age (in years). A subject-specific intercept was included as a random effect.

1 A forced-choice paradigm is ideal in signal detection designs. However, the young age of some of the subjects required that a ‘don’t know’ response option be given to all subjects. Thus, everyone was given the choice of indicating a ‘face’, ‘house’ or ‘don’t know’ response.
Results

At each presentation rate, there was a main effect of age (at 17 msec: $B = 1.93; Z = 4.89, p < .0001$; at 33 msec: $B = 1.00, Z = 3.68, p < .0002$), where percent identification improved with increasing age. There was no main effect for object type on percent identification in the presence of the interaction. There was a statistically significant interaction of age and object type on percent identification only for the 17-msec presentation rate ($B = -0.62, Z = -2.01, p < .05$; see Figure 2), where the slope for face identification became steeper than for house identification as age increased. The interaction was non-significant at the 33-msec rate ($B = 0.08, Z = 0.45, ns$).

It was possible that the youngest groups may have been more likely to say they saw nothing (i.e. ‘miss’), so post-hoc correlations were estimated on the mean number of misses between-subjects and age. The relative absence of the face advantage effect in the youngest subjects was not a reflection of more misses at younger ages. When controlling for percent identification, there was not a significant linear relationship between age and likelihood of indicating a miss, $r (84) = -0.15, p = .16$.

Although the younger participants were more likely to complete less of the experiment than older participants due to boredom or fatigue, this did not affect the ratio of face and house trials (which was equal; see Table 1). Furthermore, as can be seen in Table 1, this age difference in number of trials completed only existed for the slower presentation (33 msec), where results show there was no difference in face and house identification at any age. There was no significant correlation between age and trial number for faces ($r (84) = .18, p = .10$) or houses ($r (84) = .17, p = .12$) at the 17-msec presentation, although there was for faces ($r (80) = .47, p < .001$) and houses ($r (80) = .48, p < .001$) at the 33-msec presentation.

Discussion

Faces belong to a class of objects with which humans have extended experience. As evidenced by faster visual processing of face stimuli, humans are remarkably efficient in processing faces relative to other objects (Hoshiyama et al., 2003; Purcell & Stewart, 1988; Shelley-Tremblay & Mack, 1999). The results from the current study show a face advantage effect at rapid stimulus presentation speeds, where faces were identified more readily than houses, and the size of this advantage was shown to be age-dependent.

General identification of objects, whether houses or faces, improved gradually with age as shown by steadily improving percent identification scores. However, the slope was steeper for faces than it was for houses at the faster presentation rate. The size of the face advantage effect varied with age so that the youngest participants’ face identification scores were similar to their house identification scores, and the disparity between the identification scores was largest for the oldest participants. A post-hoc comparison showed that these changes in percent identification at 17 msec were not the result of developmental changes in response biases (i.e. increased reported ‘misses’ for the youngest groups) or the result of fewer trials completed. There was no interaction at the 33-msec presentation rate, presumably because the presentation rate was above visual thresholds and thus participants were at or near ceiling levels of identification for their age group for both classes of objects.
Our ability to recognize a highly familiar object, like a face, developmentally precedes our ability to recognize a less familiar object. These data show that faces maintain a more robust representation than houses and the robustness is age or possibly experience dependent. A working hypothesis for these findings is that early in life, there is no processing advantage for identifying any object class, although there may be a preference for certain face-like spatial arrangements. As we gain more visual experience with one class of objects, those objects may gain certain visual privileges (Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski & Gore, 1999). In other words, experience with a class of objects stabilizes our representations for that object so that subsequent encounters with it are facilitated. It is the case that experiences alter or shape visual perceptual and neural processes, as in the case of recognizing same-species faces (Pascalis, de Haan & Nelson, 2002; Pascalis, Scott, Kelly, Shannon, Nicholson, Coleman & Nelson, 2005), words (McCandliss, Cohen & Dehaene, 2003; Shaywitz, Shaywitz, Blachman, Pugh, Fulbright, Skudlarski, Menc, Constable, Holahan, Marchione, Fletcher, Lyon & Gore, 2004), non-native faces (Sangrigoli & de Schonen, 2004) and items with which we have unique expertise (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Gauthier et al., 1999).

By presenting the stimuli at speeds that lie at the boundary of perceptual recognition, face–object discriminations (or sensitivity to first-order relations) could be examined at different time points in development. As in previous studies (Purcell & Stewart, 1988; Shelley-Tremblay & Mack, 1999), it was found that faces have faster access to the nervous system. Passarotti and colleagues (2003) identified a response in the fusiform gyrus to faces in children aged 10–12 that appears to be more distributed than the typical pattern identified in adults. The area of the fusiform gyrus and the surrounding ventral occipital and temporal cortices have been implicated in face processing, and activity in this region has been suggested to be modulated by level of expertise with a particular object class (Gauthier & Tarr, 1997; Gauthier et al., 1999; McCandliss et al., 2003). Furthermore, face detection at speeds similar to those used in this study was correlated with increased activity in this ventral occipitotemporal area in adults (Grill-Spector, Knouf & Kanwisher, 2004). Therefore, late emergence of the face advantage effect may reflect the protracted development of the ventral occipitotemporal cortex response to faces throughout middle and late childhood.

There are two potential limitations in this study. First, most psychophysiological studies using backward masking paradigms control for contrast differences between object types. In this study we controlled for luminance rather than contrast differences (Allison, Puce, Spencer & McCarthy, 1999). It may be argued that the potential contrast difference between houses and faces may have contributed to the difference in performance in discriminating these object types. However, other studies (e.g. Reinders et al., 2005; Shelley-Tremblay & Mack, 1999) have found a processing advantage for faces, even when perceptual variables like contrast were equated across object types. Moreover, if it were the case that low-level perceptual differences were driving the face advantage effect (i.e. if faces were simply easier to see by virtue of contrast differences), then faces should be better detected at all ages, not just for the older subjects. Secondly, in the current study, faces were only compared to one object category, houses. Ideally, as suggested by Gauthier and Nelson (2001), studies should make use of face–object–object comparisons. Without more categories, interpretations about face processing may be incomplete and overestimate the uniqueness of face processing skills. Future studies should examine the face advantage in the context of a face–object–object design.

Overall these findings suggest that face processing skills can be assessed across ages with rapid presentation paradigms. Methodologically, this study shows that when studying face versus non-face object processing using masking techniques, presentation times must be catered to the age of the subject and to the stimulus type, since amount of time needed to process images varies as a function of these variables. While this study cannot address whether the face advantage effect exists as a result of extended experience with faces or not, it does establish that face representations become more robust with increasing age throughout childhood and adolescence. The finding that both face and house processing slopes are gradual and linear, but that the slope for face processing is steeper than for houses is consistent with the notion that large amounts of experience with an object can speed up the development of the

---

Table 1  Number of trials completed in each condition split by age group; mean (SD)

<table>
<thead>
<tr>
<th></th>
<th>Face 17 msec</th>
<th>House 17 msec</th>
<th>Face 33 msec</th>
<th>House 33 msec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children (5–12)</td>
<td>34.71 (15.53)</td>
<td>34.69 (14.66)</td>
<td>17.62 (10.13)</td>
<td>16.82 (9.75)</td>
</tr>
<tr>
<td>Adolescents (13–17)</td>
<td>36.44 (11.84)</td>
<td>37.50 (10.98)</td>
<td>22.06 (8.71)</td>
<td>21.11 (9.39)</td>
</tr>
<tr>
<td>Adults (18–35)</td>
<td>40.00 (11.61)</td>
<td>39.76 (10.68)</td>
<td>29.48 (11.82)</td>
<td>29.40 (12.01)</td>
</tr>
</tbody>
</table>

© 2006 The Authors. Journal compilation © 2006 Blackwell Publishing Ltd.
processing skill for that object. While this issue can only be addressed with training studies, these data suggest that our massive amounts of experience with faces speeds up the development of face processing skills. Thus, across development, faces are not a special class of objects, but instead become special to us because of their high frequency.

References


Received: 8 July 2004
Accepted: 13 September 2005