Infant categorization of faces: Ladies first

Jennifer L. Ramsey a,*, Judith H. Langlois a, Nathan C. Marti b

a Department of Psychology, The University of Texas at Austin, USA
b The Community College Survey of Student Engagement, The University of Texas at Austin, USA

Received 14 October 2004; revised 21 December 2004
Available online 1 April 2005

Abstract

We review and provide empirical evidence to show that infants categorize and process male and female faces differently, with an advantage in processing female faces. To understand this asymmetry in categorization and processing of male and female faces, we evaluate three mechanisms influencing infant categorization of male and female faces: differential experience with female and male faces; early visual preferences for female vs. male faces; and range of physical differences among category exemplars. The paper concludes with a developmental trajectory for infant acquisition of face categories proposed within a framework that reflects current knowledge and theory in the infant categorization and face processing literatures. The proposed developments have important implications for the existing infant face perception literature and infant learning about females and males.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Infant development; Face perception; Categorization; Sex differences; Facial prototype; Social development; Experience; Visual preferences; Variability

* A portion of this work was previously presented at the May 2004 International Conference on Infant Studies in Chicago, IL. This work was supported by a grant to Judith H. Langlois from the National Institute of Child Health and Human Development.
* Corresponding author. Present address: Department of Psychology, University of Nevada Las Vegas, USA. Fax: +1 702 895 0195.
E-mail address: ramseyj2@unlv.nevada.edu (J.L. Ramsey).
Do infants process and categorize male and female faces similarly? Although a few studies suggest infants categorize both female and male faces naturally and effortlessly (Cornell, 1974; Leinbach & Fagot, 1993; Younger & Fearing, 1999), this paper will review literature and provide empirical evidence showing a substantial developmental asymmetry in infants’ processing and categorization of male and female faces with male faces being significantly more difficult for young infants. We then review research to suggest three reasons underlying this asymmetry. Last, we propose a developmental trajectory for infant learning of faces and subsequent grouping into male and female. Because developmental differences in infants’ learning of male vs. female faces may affect conclusions from the existing face perception literature, the paper concludes with caveats for generalizing results from infant studies using female faces to all faces.

Categorization

Categorization of natural objects is a fundamental cognitive activity for all human beings (Mervis & Rosch, 1981). Cognitive categorization guides the grouping of objects and events in the world into different classes. Category members are distinguishable, but share particular characteristics, and individuals treat or act upon these objects or events similarly (e.g., Mervis & Pani, 1980; Mervis & Rosch, 1981). Through cognitive categorization, people can, on average, consistently respond to novel members of a category based upon their knowledge of that category. Although such simplification of the world results in some information loss, the value of predicting the utility or behavior of a particular object based on prior experience with members of that category would have been adaptive for the survival of ancestral humans (e.g., Humphrey, 1980; Quinn & Eimas, 1987).

Given the importance of categorization, it is not surprising that even young infants demonstrate this ability (e.g., Eimas, 1994; Mandler, 1992; Mervis & Rosch, 1981; Quinn & Eimas, 1987, 1996). Infants are equipped with a predisposition to develop a system of categories given appropriate experience (e.g., Behl-Chadha, 1996; Cohen & Strauss, 1979; Eimas & Quinn, 1994; Humphrey, 1980; Sherman, 1985). Infants categorize both natural objects (e.g., cats and dogs; Quinn, Eimas, & Rosenkrantz, 1993) and artificial objects (e.g., dot patterns; Younger & Gotlieb, 1988) within a laboratory setting after only minimal exposure to exemplars. Natural categories form as a result of exposure to exemplars in the regular environment. In contrast, artificial categories form as a result of exposure to exemplars within an experimental setting.

Infant knowledge of female and male face categories

One presumably natural category that infants learn is the distinction between male and female faces. Although some studies suggest that infants readily categorize both male and female faces (e.g., Cornell, 1974; Leinbach & Fagot, 1993; Younger & Fearing, 1999), we argue that the data from categorization studies more accurately
portray an asymmetry in infant categorization of faces: Infants are more proficient at categorizing and processing female than male faces. The data we review come from four different types of categorization studies: prototype formation studies; prototype preference studies; intermodal matching studies; and familiarization/habituation studies. We conclude this section with data from discrimination studies suggesting some early precursors to this asymmetry in categorization.

**Male and female facial prototypes**

Our original interest in the asymmetry in infant categorization of female and male faces stemmed from a series of studies investigating infants’ ability to form prototypes of female and male faces conducted in The Laboratory for Social Development at The University of Texas at Austin (http://www.psy.utexas.edu/LangloisLab). Results from these studies suggest that 6-month-olds can form a prototype of female faces (Rubenstein, Kalakanis, & Langlois, 1999), but not male faces (Langlois et al., 1997). The ability to form male and female face categories not only implies that infants can recognize that faces of the same sex belong to a category, but also that infants should be able to form a prototype of that category (e.g., Sherman, 1985; Strauss, 1979; Younger, 1990; Younger & Gotlieb, 1988). A prototype is an abstraction from the category exemplars that represents the category average (Rosch, 1978; Rosch, Simpson, & Miller, 1976). The ability to form a prototype implies that infants have had sufficient experience with a particular category and can therefore compare new objects to the prototype to determine whether or not they belong to that category.

To test the hypothesis that infant preferences for attractive faces may be due to their ability to average across faces to form prototypes (Langlois & Roggman, 1990), Langlois and her students tested 6-month-olds’ ability to form averages of female faces (Rubenstein et al., 1999). Infants viewed eight different individual female faces twice during familiarization trials. During test trials, infants viewed the following face pairs: a familiar face paired with a novel face; a familiar face paired with an averaged face created by digitizing the images of the eight familiar faces and mathematically averaging them together; and a novel face paired with the averaged face. The test faces were equally attractive, so infant looking time should be based on novelty, not on attractiveness. According to cognitive information processing theory, evidence that cognitive averaging of the exemplars has occurred should cause the infants to treat the averaged face as more familiar than a familiar face to which they had previously been exposed (e.g., Quinn et al., 1993). Based on the assumption that infants prefer novelty, the researchers predicted that infants would look longest at the novel face, followed by the familiar face, and then the averaged face. Indeed, they found this exact pattern of looking (Rubenstein et al., 1999). This finding has been

---

1 Although we acknowledge that there is a debate in the literature regarding whether categories are represented by exemplars or prototypes, the majority of categorization and face processing research with both infants and adults provides stronger support for prototype than exemplar models or support that both forms of representation may be present and used in different situations (e.g., Busey & Tuminclif, 1999; Cabeza, Bruce, Kato, & Oda, 1999; Leopold, O’Toole, Vetter, & Blanz, 2001; Malt, 1989; Sherman, 1996; Sherman, 1985; Smith, 2002; Strauss, 1979; Valentine & Bruce, 1986; Younger, 1990).
replicated with female faces by other labs with infants aged 3 months (de Haan, Johnson, Maurer, & Perrett, 2001) and newborns (Walton & Bower, 1993).

When, however, Langlois and students (1997) extended these studies to male faces using the same familiarization and testing procedures, they did not find that infants could average across male faces. Because infants have less experience with male than with female faces (e.g., Bailey, 1994; Harrison & Magill-Evans, 1996; Hossain & Rooopmarine, 1994), this differential experience could make the task more difficult for infants and could thus influence their familiarity/novelty preferences (Hunter & Ames, 1988). To test this possibility, Langlois et al. (1997) increased the number of visits and familiarization trials, so that infants came to the laboratory twice and viewed each of the eight male faces five times prior to participating in the test trials of the prototype formation study. Surprisingly, infants still showed no evidence of averaging across the faces. Perhaps experience with male faces outside of the experiment was necessary before infants could abstract the averaged face. Accordingly, Langlois and students (1997) tested older infants (8-month-olds), but again found no evidence of infants’ ability to average across male faces. Additional changes to the methodology, such as introducing a delay of 15 min or 1–2 days between the familiarization trials and the test trials to allow more time for the individual exemplars to fade and the averaged face to form (Homa & Vosburgh, 1976) were equally unsuccessful (Langlois et al., 1997). See Table 1 for a brief summary of the studies and their results.

Even though null results are difficult to interpret, the persistent finding that infants did not respond as predicted to the test faces during the male face prototype studies suggests that 6- and 8-month-olds cannot form a prototype of male faces. In contrast, the evidence is robust that newborns, 3- and 6-month-olds can form a prototype of female faces (de Haan et al., 2001; Rubenstein et al., 1999; Walton & Bower, 1993). Although it is unclear why these differences occurred within a laboratory setting in which infants saw a finite number of exemplars from the category, it may be that the infants were still trying to learn about male faces causing them to process the male faces more slowly than the female faces. Faces are commonly experienced natural stimuli, but are also very complicated objects. Therefore, when the researchers presented infants with exemplars of male faces in the laboratory setting, infants may have used this opportunity to acquire knowledge about them. With familiar female faces, in contrast, infants could proceed to form more complex representations. Thus, we decided to investigate if this asymmetry in infant categorization of female and male faces appeared in other studies.

Preferences for facial prototypes

Because attractive faces are prototypical faces (Langlois & Roggman, 1990; Rhodes, Sumich, & Byatt, 1999; Rhodes & Tremewan, 1996), and because people tend to prefer the prototype of a category (e.g., Halberstadt & Rhodes, 2000; Piter & Stokmans, 2000; Whitfield & Slatter, 1979), the ability of infants to form prototypes of female faces has been used as an explanation for their preferences for attractive female faces (de Haan et al., 2001; Rubenstein et al., 1999). Thus, further support for infants’ ability to form prototypes of female faces stems from their preferences for attractive female faces. If 6-month-old infants are unable to form prototypes of male
Table 1
Summary of studies assessing infants’ ability to form prototypes of male face

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Description of familiarization trials</th>
<th>Test trial comparisons</th>
<th>Mean looking times</th>
<th>Significance (only p values &lt;.10 are listed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>8 individual male faces seen 2 times PIP; 5-s trials</td>
<td>Composite vs. novel</td>
<td>4.96 vs. 5.66</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. novel</td>
<td>6.21 vs. 4.50</td>
<td>p = .08</td>
</tr>
<tr>
<td>9</td>
<td>8 individual male faces seen 3 times PIP; 5-s trials</td>
<td>Composite vs. novel</td>
<td>6.59 vs. 7.04</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. novel</td>
<td>6.69 vs. 6.49</td>
<td>n.s.</td>
</tr>
<tr>
<td>63</td>
<td>8 individual male faces seen 1 time; infant-controlled trials</td>
<td>Composite vs. novel</td>
<td>4.53 vs. 4.77</td>
<td>n.s.</td>
</tr>
<tr>
<td>80</td>
<td>8 individual male faces seen 2 times PIP; 5-s trials</td>
<td>Composite vs. novel</td>
<td>5.56 vs. 6.05</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. novel</td>
<td>6.21 vs. 5.33</td>
<td>p = .09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. composite</td>
<td>6.41 vs. 5.25</td>
<td>p = .01</td>
</tr>
<tr>
<td>54</td>
<td>8 individual male faces seen 2 times PIP; 5-s trials</td>
<td>Composite vs. novel</td>
<td>4.96 vs. 4.89</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. novel</td>
<td>5.65 vs. 4.44</td>
<td>p = .01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. composite</td>
<td>5.37 vs. 4.89</td>
<td>n.s.</td>
</tr>
<tr>
<td>30</td>
<td>8 individual male faces seen 2 times PIP; 5-s trials</td>
<td>Composite vs. novel</td>
<td>5.13 vs. 4.19</td>
<td>p = .08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. novel</td>
<td>4.95 vs. 3.80</td>
<td>p = .02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Familiar vs. composite</td>
<td>5.02 vs. 4.58</td>
<td>n.s.</td>
</tr>
<tr>
<td>N</td>
<td>Description</td>
<td>Composite vs. novel</td>
<td>Familiar vs. novel</td>
<td>Familiar vs. composite</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>21</td>
<td>8 individual male faces seen 2 times PIP; 5-s trials; 15-min break between familiarization and test trials</td>
<td>4.49 vs. 4.37</td>
<td>5.29 vs. 3.96</td>
<td>5.41 vs. 4.68</td>
</tr>
<tr>
<td>25</td>
<td>8 individual male faces seen 2 times PIP one day and then seen 2 times PIP 1–3 days later; 15-min break; same 8 individual male faces seen 1 more time PIP; 5-s trials</td>
<td>4.78 vs. 4.04</td>
<td>4.61 vs. 4.63</td>
<td>4.24 vs. 3.76</td>
</tr>
<tr>
<td>33</td>
<td>8 individual male faces seen 2 times PIP; 15-min break between familiarization and test trials</td>
<td>4.35 vs. 4.56</td>
<td>5.10 vs. 4.75</td>
<td>5.06 vs. 4.36</td>
</tr>
<tr>
<td>56</td>
<td>8 individual male faces seen 2 times PIP; 1–2 day break between familiarization and test trials</td>
<td>4.53 vs. 4.18</td>
<td>4.57 vs. 4.49</td>
<td>5.53 vs. 4.46</td>
</tr>
<tr>
<td>21</td>
<td>8 individual male faces seen 2 times PIP; 1–2 day break between familiarization and test trials</td>
<td>5.06 vs. 4.81</td>
<td>5.20 vs. 4.37</td>
<td>6.37 vs. 4.19</td>
</tr>
</tbody>
</table>

**Note.** PIP means presented in pairs. All infants saw the test face pairs twice during 10-s trials with the left–right positions reversed during the second showing. Mean looking times are in seconds.
faces, then they should also not show preferences for prototypes of male faces, or attractive male faces. Four different groups of researchers have investigated infant preferences for attractive male faces and the results are somewhat mixed (Kramer, Zebrowitz, Giovanni, & Sherak, 1995; Langlois, Ritter, Roggman, & Vaughn, 1991; Ramsey, 2003; Samuels & Ewy, 1985). In two studies, 6-month-olds and 3- and 6-month-olds, respectively, saw pairs of high and low attractive faces with pairs of female and male faces seen on alternating trials (Langlois et al., 1991; Samuels & Ewy, 1985). When the researchers analyzed the looking time data to the high and low attractive faces, they collapsed the data across both the female and male faces and found a significant preference for the attractive faces with no sex of stimulus interaction, suggesting infants have preferences for both attractive female and male faces. In another study, however, 4.5- to 6-month-olds also saw alternating pairs of high and low attractive female and male faces, matched on babyfaceness, and the results showed a sex of stimulus interaction: Infants showed a visual preference for the attractive female faces, but not the attractive male faces (Kramer et al., 1995). Thus, the results for studies in which infants saw both female and male face pairs provide mixed results regarding infants’ visual preferences for attractive male faces.

Because seeing female faces may “prime” infant interest in attractive male faces (Langlois et al., 1991), it is unclear if the significant results from the above studies would replicate if infants saw only male face pairs during the study. In addition, infants’ lack of preference for attractive male faces was found in a study that held babyfaceness constant within the face pairs (Kramer et al., 1995). Thus, it may be important to hold constant certain cues that may affect infant interest in male faces when assessing attractiveness preferences. The masculinity/femininity of male faces affects adult interest in male faces (e.g., Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000; Thornhill & Gangestad, 1993), and therefore may also influence infant interest in male faces. To determine the robustness of infant preferences for attractive male faces when they see only male faces, Ramsey (2003) showed 6- and 12-month-olds pairs of high and low attractive male faces, that either matched in masculinity (Study 1) or matched or differed in masculinity (Study 2). She found no consistent preferences for high attractive male faces in general, despite using different face pairs in two separate studies in which the sample size was large enough to detect an effect if one existed (Ns ranged from 54 to 69 for each particular age group within the two studies; Buchner, Faul, & Erdfelder, 1997). Because several studies have demonstrated infant preferences for attractive female faces (Kramer et al., 1995; Langlois et al., 1987, 1991; Samuels & Ewy, 1985; Samuels, Butterworth, Roberts, Graupner, & Hole, 1994; Slater, Bremmer et al., 2000; Slater, Quinn, Hayes, & Brown, 2000; Slater et al., 1998), but the findings for

---

2 In Study 1, Ramsey (2003) found a main effect for attractiveness (suggesting visual preferences for low attractive male faces) that was superceded by an interaction between attractiveness of the faces and masculinity of the face pair: Infants showed visual preferences for high attractive faces when the face pair was low masculine, visual preferences for low attractive faces when the face pair was medium masculine, and no significant visual preferences for attractiveness when the face pair was high masculine. In Study 2, there was no main effect or interaction with attractiveness.
attractive male faces are mixed and therefore not as robust, it appears young infants have consistent preferences for prototypical female faces, but not prototypical male faces. This difference in infant preferences is likely the result of infants not yet having a prototype for male faces because their category of male faces is not as well developed as their category of female faces. Both the prototype formation and prototype preference studies suggest that infants may have a more mature category for female than male faces, so we reviewed other categorization studies to determine if this asymmetry appeared in other data.

Intermodal matching studies

Another method used to assess infant knowledge of male and female face categories is to investigate their intermodal knowledge of female and male faces and voices (Patterson & Werker, 2002; Poulin-Dubois, Serbin, & Derbyshire, 1998; Poulin-Dubois, Serbin, Kenyon, & Derbyshire, 1994; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991). Intermodal knowledge refers to the ability to match things that go together across different sensory modalities (e.g., matching a female face with a female voice or a male face with a male voice). This method is a useful way to determine infants’ development of male and female face categories in the natural environment because infants need to use their many different experiences with male and female faces to demonstrate actual knowledge of the category (Reznick, 1989). Although this method requires more attention per se than methods that require only attention to a face or voice, there is much evidence that infants use and link together information from across different sensory modalities at a very early age and that infants are more likely to learn information via a combination of sensory pathways rather than through a single sensory pathway (the intersensory redundancy hypothesis; Bahrick & Lickliter, 2000, 2002).

Studies using intermodal matching to assess infants’ knowledge of female and male faces and voices have found varying results. Whereas Walker-Andrews et al. (1991) found that infants as young as 6 months of age could match a female face with a female voice and a male face with a male voice, Patterson and Werker (2002) found that although infants showed improvement in matching faces and voices by sex around 6 months of age, it was not until they were 8 months of age that they could reliably match the appropriate face and voice. In contrast, Poulin-Dubois et al. (1994) found that 9- and 12-month-olds could match female faces and voices, but the 9-month-olds could match only very stereotypical female faces and voices, and neither age group could match male faces and voices. A follow-up study found that infants could not match male faces and voices until 18 months of age (Poulin-Dubois et al., 1998). Even at this age, however, the means for the matching and mismatching data suggest the infants were more accurate at matching female than male stimuli. Whereas the first two studies suggest that infants can match the faces and voices of both sexes between 6 and 8 months of age (Patterson & Werker, 2002; Walker-Andrews et al., 1991), the second two studies suggest that infants can match only female faces and voices around 9 months of age, but cannot match male faces and voices until 18 months of age (Poulin-Dubois et al., 1994, 1998).
These discrepant results may have occurred because of differences in stimuli. For the first two studies, infants saw videotaped, moving images of a female and a male face speaking while a voice played (Patterson & Werker, 2002; Walker-Andrews et al., 1991), whereas for the second two studies, infants saw static images of female and male faces while a voice played (Poulin-Dubois et al., 1994, 1998). Using dynamic, rather than static, stimuli in the first two studies may have made the task easier because of its similarity to real-life situations and may therefore be one reason why younger infants demonstrated intermodal knowledge of females and males.

The first two studies, however, were not necessarily a test of categorical knowledge because infants saw only one pair of female and male faces (Patterson & Werker, 2002; Walker-Andrews et al., 1991). These faces were fairly stereotypical, at least in the Patterson and Werker (2002) study; Walker-Andrews et al. (1991) did not describe the stereotypicality of their stimuli. For the second two studies, infants saw several pairs of female and male faces that ranged in stereotypicality (Poulin-Dubois et al., 1994, 1998). These latter two studies more likely reflected a test of categorical knowledge than the studies using only one pair of male and female faces because of the range and number of male and female faces included. Any test of categorical knowledge should include several exemplars that are characteristic of that category and not just “good” or prototypic examples of the category. Hence, the results from intermodal studies testing categorical knowledge suggest an asymmetry in infant categorization of females and males.

Taken together, the results from the prototype formation, prototype preference, and intermodal matching studies point to an asymmetry in the development of infants’ categorical knowledge of female and male faces, with their category of female faces being more advanced in development than their category of male faces. Performance within these studies requires that infants bring real-world knowledge into the laboratory setting. Infants need to have a prototype of male faces before they can exhibit a preference for prototypic (i.e., attractive) male faces and it appears that infants have difficulty forming prototypes of male faces even within a laboratory setting with exposure to a limited number of faces. Infants also need to have enough experience with males to match a series of their faces and voices together in a laboratory setting. In the final series of categorization studies to be reviewed, familiarization/habituation studies, we found not only evidence of an asymmetry in infant categorization of female and male faces, but also that the literature is lacking strong evidence that infants can categorize the sex of faces within the normal range of the population of faces.

**Familiarization/habituation studies**

Several researchers assessed infants’ ability to group male and female face categories using familiarization/habituation techniques (Cornell, 1974; Leinbach & Fagot, 1993; Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002; Younger & Fearing, 1999). In one study, researchers familiarized 4.5- to 6-month-olds to a series of female or male faces and then tested them with a novel face from the familiar sex category and a novel face from the other (novel) sex category (Cornell, 1974). Based on infants’ novelty preferences, Cornell (1974) concluded that 5.5- to 6-month-olds demonstrated an
ability to categorize faces as male or female. Cornell (1974), however, analyzed the data for infants’ ability to categorize sex of face (Condition 1) with two other conditions assessing infants’ ability to recognize individual faces from differing orientations (Condition 2) and infants’ ability to recognize individual faces from only one orientation (Condition 3). Because the results from each condition were analyzed together, rather than separately, and infants’ recovery to the novel stimulus was greatest during Condition 3, it is unclear whether the results from the sex categorization condition were significant on their own. It is also unclear if infants could discriminate among the faces used in the study because there were no tests of within-category face discrimination (Quinn et al., 2002). Thus, we do not know whether the infants were treating faces within categories as individual faces belonging to the same category or treating the within-category faces as the same face, making the results ambiguous (Quinn, 1987). These results, therefore, do not provide strong support for the conclusion that infants categorize faces by sex.

In a study using a similar familiarization and testing paradigm with infants aged 5, 7, 9, and 12 months, only the older infants demonstrated evidence of an ability to categorize by sex—the majority of the 5- and 7-month-olds did not dishabituate to the novel face of the other sex during the test trials (Leinbach & Fagot, 1993). Although the results from this study suggest that infants develop the ability to reliably categorize male and female faces around 9 months of age, this conclusion is problematic for a few reasons. First, the 12-month-olds’ performance significantly decreased when the researchers altered cues such as clothing and hairstyle to make the males and females look more unisex (Leinbach & Fagot, 1993), suggesting that the categorization is highly dependent on cues other than facial structure, even at this older age. Therefore, the infants may have been categorizing the faces based on hair cues (i.e., long- and short-haired people) rather than based on sex as conveyed by the face (Fagot & Leinbach, 1993).

In addition, the faces used in the study were from magazines and catalogs and were highly attractive and sex-stereotypical and not necessarily representative of female and male faces infants usually see within their natural environments. Because attractive faces are more typical-looking (i.e., not unusual) than less attractive faces (Langlois & Roggman, 1990; O’Toole, Deffenbacher, Valentin, & Abdi, 1994; Vokey & Read, 1992), there is less variability among attractive faces, suggesting that the faces used by Leinbach and Fagot (1993) were more similar to one another than faces of varying levels of attractiveness would have been. This similarity within categories may have made categorization easier for the infants (Mervis & Rosch, 1981; Oakes & Spalding, 1997; Quinn et al., 1993). Indeed, recent research suggests that infants as old as 11 months of age have difficulty categorizing the sex of faces if the stimulus faces are less attractive, less stereotypical looking, and therefore more dissimilar (Newell & Strauss, 2002; Newell, Strauss, & Best, 2003). Thus, it is unclear how generalizable Leinbach and Fagot’s (1993) results are to the normal population of male and female faces.

Finally, the infants in this study showed greater recovery to a female face after seeing male faces during the habituation trials than they did to a male face after seeing female faces during the habituation trials (Leinbach & Fagot, 1993). This greater
recovery to female than male faces after equal experience with faces of the other sex during the familiarization trials suggests greater visual interest in female than male faces. Greater infant attention to female compared with male faces is not limited to this particular study. Younger infants, aged 3 to 4 months, looked longer at a novel female than a novel male face following familiarization to a series of male faces, but looked equally at a novel female and a novel male face following familiarization to a series of female faces (Quinn et al., 2002). Prompted by the results from both their study and Leinbach and Fagot’s (1993) study, Quinn et al. (2002) tested this asymmetry in infant categorization in another study that demonstrated that infants show visual preferences for female relative to male faces. This asymmetry in visual preferences suggests that it is difficult to assess infant categorization of female and male faces using familiarization/habituation studies. Longer looking to one face over the other during the test trials may occur because of a priori visual preferences or because of infants’ experience with the faces seen during the familiarization/habituation trials. If researchers want to use such methods, they need to first assess visual preferences for the test faces used and then adjust the data analysis accordingly to account for any significant differences in visual looking.

In another study, 7- and 10-month-olds saw faces of both sexes rather than just one sex during the familiarization trials—they saw typical-looking female and male faces randomly presented with the constraint that they could not see more than three faces of one sex presented in a row (Younger & Fearing, 1999). During the test trials, infants saw a novel sex-typical face and a novel sex-ambiguous face (a face with which adults could not easily judge the sex). The 7-month-olds did not dishabituate to either novel face, suggesting that they were classifying male and female faces into a single category (i.e., faces in general) rather than into separate sex categories. In contrast, the 10-month-olds did dishabituate to the sex-ambiguous faces. The researchers interpreted this result as evidence that the 10-month-olds were treating the familiarization faces as faces from two separate categories (male and female) and that the sex-ambiguous test faces did not easily fit into one of those two categories. Although the authors did not find an interaction with sex of stimulus for this effect, the means suggest that infant dishabituation to the sex-ambiguous female face was much stronger than dishabituation to the sex-ambiguous male face. According to the log-transformed looking times, infant looking at the sex-ambiguous vs. the sex-typical faces was 1.12 vs. .74 for the female test faces and .80 and .73 for the male test faces (Younger & Fearing, 1999). Even though the looking toward the sex-ambiguous male test face was in the correct direction, thus eliminating a sex of stimulus interaction, it is likely that the looking toward the sex-ambiguous female test face carried the significant results of the 10-month-olds in this study. The results therefore seem to provide evidence that infants have a category for female faces, albeit typical-looking female faces, but not necessarily as complete of a category for male faces.

A further complication applies to all of the above-mentioned familiarization/habituation studies. Evidence that infants categorize a particular group of objects within the laboratory does not necessarily imply that they possess knowledge of that category within their natural environment. Infants may have learned the category during the experimental familiarization process (Reznick, 1989), suggesting that just
because infants can form a category in an artificial experimental situation does not mean that infants do make the same categorical distinctions in their natural environment. Thus, the extant literature using familiarization/habituation procedures to test infant categorization of sex of face has not provided strong evidence that infants can categorize the sex of faces seen within the range of their natural environment. In addition, the discrepancies in infant recovery to female and male faces during the test trials also suggest an asymmetry in infant categorization of female and male faces that may originate from discrepancies in infant discrimination/recognition abilities.

**Discrimination/recognition studies**

A prerequisite for forming a category for faces in general or separate categories for female and male faces is that infants must discriminate faces within each category and between the two face categories. If infants cannot discriminate faces, they may perceive all within-category exemplars as the same face (Quinn, 1987). When such infants treat a novel example of the novel category as different from a novel example of the familiar category, they are not categorizing, but failing to discriminate highly complex faces.

Early in development, asymmetry in infant discrimination within female and male face categories can be found. Infants visually discriminate their mother’s face from a female stranger’s face as young as 2-days-old (Bushnell, Sai, & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995; Walton, Bower, & Bower, 1992), but do not show this same difference in looking upon seeing their father’s and a male stranger’s face at this age (Walton et al., 1992) or even at 4 months of age (although some evidence indicates a difference in 4-month-olds’ affective responses toward their father’s and a male stranger’s face; Ward, 1998). In addition, if given sufficient familiarization to a stranger’s face, infants can tell the difference between two similar-looking female strangers as early as 3 months of age (Barrera & Maurer, 1981a). At an older age, however, 7- to 7.5-month-olds have difficulty discriminating among adult-judged, similar-looking male faces (although they can discriminate among adult-judged, highly discriminable male faces; Fagan, 1976). This early asymmetry in discrimination abilities is reflected in 3- to 4-month-olds’ ability to recognize a previously seen exemplar from a series of female faces, and their inability to recognize a previously seen exemplar from a series of male faces (Quinn et al., 2002).

Although not all studies find the same age at which infants show between-category discrimination of males and females (most likely due to differences in procedure and stimuli), sometime between 5 and 8 months infants show between-category discrimination of male and female faces (Fagan, 1972, 1974, 1976; Fagan & Singer, 1979; Lasky, Klein, & Martinez, 1974; Yamaguchi, 2000). Despite these between-category discrimination abilities, the within-category discrepancies seen earlier in development suggest an early asymmetry in infant discrimination of female and male faces, an important precursor to infant categorization of faces by sex. Quinn et al.’s (2002) work also suggests that before investigating infant categorization of female and male faces, it may be important to demonstrate not only that infants can discriminate among category exemplars, but that they can recognize individual category exem-
pals after exposure to a series of faces. Face recognition after familiarization to a series of faces is a more difficult task than face recognition following familiarization to one face because it requires not only perceiving that the faces are different from one another, but having enough familiarity with the category to notice and process slight differences among the faces so as to remember them accurately.

In sum, the infant discrimination literature suggests precursors to the asymmetry seen in categorization studies. The asymmetry suggests infants are not as expert at processing male faces as they are at processing female faces. Thus, it may be that they use exposure to male faces in laboratory settings to learn about male faces rather than perform the intended experimental task.

**Face perception tasks affected by ill-defined male face categories**

If infants use their experience in the laboratory to learn about male faces, they should look longer in studies using male faces as stimuli than in studies using female faces as stimuli. Such a discrepancy in looking times to male and female faces is likely to increase as the task becomes more complex (Hunter & Ames, 1988), particularly when the task requires high information processing demands such as categorical abstraction (e.g., recognizing which members do and do not belong to a particular category or forming a prototype or average of the category).

To examine the hypothesis that infants look longer in tasks using male relative to female faces, we conducted a meta-analysis of 15 infant face perception studies \((N = 1069 \text{ six-month-olds})\) from The Laboratory for Social Development at The University of Texas at Austin (http://www.psy.utexas.edu/LangloisLab). The studies differed in level of complexity, sex of stimulus faces, and type of procedure; therefore, we coded for these study characteristics. For example, some of the studies were preferential looking tasks conducted to assess infant interest in certain types of faces (e.g., attractive vs. unattractive faces). We coded preference studies as low in complexity because infants needed only to look at the faces and reveal their natural preferences, requiring nominal information processing. Other studies were discrimination tasks assessing infants’ ability to discriminate between individual faces from the same category (e.g., recognizing a male face as different from a previously shown male face). We coded these studies as moderate in complexity because the task required recognition memory of previously shown faces. Finally, some studies were categorization tasks that required infants to either group faces into certain categories (attractive vs. unattractive) or to form a prototype from exposure to exemplars of faces. We coded these studies as high in complexity because the task not only required recognition, but some form of categorical processing. Some infants participated in studies with only female or only male faces, whereas other infants participated in studies with both sexes shown, but on different trials. Thus, we coded for sex of stimulus face. Finally, we coded for type of procedure: infant-controlled, in which the faces stayed up on screen as long as the infants looked at them; or fixed trial length, in which the faces stayed up on screen for a predetermined amount of time.

Meta-analysis typically uses standardized effect size values, such as an \(r\) or \(z\), as a dependent variable in random effects models. Because we had complete data, however,
it was not necessary to estimate variance based on the sample size or to use aggregate statistics such as an $r$ or $z$. Rather, we used random intercepts models in the hierarchical linear modeling (HLM) framework with complete within-study data (Raudenbush & Bryk, 2002). HLM modeling separates variation from individuals and the larger units to which they belong, which in this case were individual studies. Separating these variance components was essential because there were several factors that were not controlled across the individual studies in this analysis and it was possible that variables of interest could be confounded with types of studies. Furthermore, the independent variables were not balanced because the analysis made use of existing data across several studies and thus did not adhere to an experimental design.

Ideally, each individual study would contain both male and female faces as stimuli and the difference in looking times could be used to compute a standardized effect representing the difference between looking times to male and female faces. Of the studies available, however, most used faces of a single sex, so we could not compute a standardized effect. Using the raw looking time was not a viable option either because there was a wide range in the average looking times across studies because of differences in experimental methods (some procedures were infant-controlled, some were fixed and the number of trials in each study differed). Therefore, we computed each infant’s total looking time toward the faces used in the study and divided the total by the number of trials to create an average looking time per trial. For studies in which infants saw both sexes, we computed separate totals and averages for each sex of face based on the trials infants saw these faces. To diminish the effect of the different experimental methods, we standardized looking times around type of study (infant-controlled vs. fixed) and complexity (low, medium, and high). Preliminary HLM models indicated that this procedure successfully removed the effects of type of study and complexity; neither of these variables was a significant predictor of the outcome with data in standardized form in contrast to the effects of these variables in their unstandardized form. There was still a significant random effect for study, which reflects variance accounted for by the individual studies. To separate this variability from the theoretically interesting fixed effects in the model, we used random intercepts (Raudenbush & Bryk, 2002).

Using the sex of the face stimuli as a level-1 independent variable, the HLM model showed that there was a significant effect for looking time: Infants looked longer at male compared to female faces across the studies in the model, $t(1067) = 2.04$, $p < .05$; see Fig. 1. The sex of stimulus face was treated as a fixed parameter because there was not a significant random effect and model convergence was poor when it was treated as such, thus indicating that the effect for sex of stimulus was constant across studies.

The standardized and raw values indicate that the discrepancy in looking at male relative to female faces differs depending on the complexity of the study (see Figs. 1 and 2). The pattern suggests that as studies increase in complexity, infants look longer at male compared to female faces. Because we standardized looking times around task complexity, it was not possible to directly test effects and interactions of task complexity in the current investigation, but the pattern of results suggests that an interaction of sex of stimulus and complexity of task on infant looking times is worth systematically exploring in future experimental studies.
Our meta-analysis shows that infants take more time to process male than female faces. This observation, when coupled with the finding that infants cannot perform more complex tasks with male faces (e.g., averaging across male faces), but can perform such tasks with female faces, suggests that studies using male faces are difficult for infants, particularly highly complex, categorical processing tasks. As a result, infants may use exposure to male faces in the laboratory as an opportunity to acquire more information about male faces instead of performing the experimenter’s intended task.

**Importance of problem**

Taken together, the results from the reviewed categorization studies and those from our meta-analysis suggest an asymmetry in how infants categorize and process female
and male faces. Specifically, infants demonstrate a greater advantage in categorizing and processing female faces. Understanding how infant categorization and processing of female and male faces might differ is important because the infant face perception literature is replete with studies investigating infant interest in faces, face recognition, preferences for certain types of faces, understanding of emotion, and understanding of social interaction through facial cues (e.g., Barrera & Maurer, 1981b; Caron, Caron, & MacLean, 1988; de Haan & Nelson, 1999; Fagan & Singer, 1979; Field et al., 1984; Hains & Muir, 1996a, 1996b; Kestenbaum & Nelson, 1990; Langlois, Roggman, & Rieser-Danner, 1990; Ludemann & Nelson, 1988; Nelson & Collins, 1992; Pascalis, de Haan, Nelson, & de Schonen, 1998; Walker-Andrews, 1986; Weinberg & Tronick, 1996). Very few, less than 20% of the studies we reviewed, however, included both male and female faces as stimuli (Braungart-Rieker, Garwood, Powers, & Notaro, 1998; Cornell, 1974; Dirks & Gibson, 1977; Fagan, 1972; Fagan & Singer, 1979; Haith, Bergman, & Moore, 1977; Langlois et al., 1991; Maurer & Salapatek, 1976; Nelson & Dolgin, 1985; Phillips, Wagner, Fells, & Lynch, 1990; Samuels & Ewy, 1985). Indeed, the majority of studies made use of only female faces. Many primary researchers and interpreters of primary research seem to presume that, as revealed by statements about infant “face” perception, their results with female faces generalize to all faces. We believe this generalization is an important oversight.

In addition, there has been a recent focus on asymmetries seen in both infant face processing and categorization (e.g., Nelson, 1993; Pascalis, de Haan, & Nelson, 2002; Quinn, 2002; Sangrigoli & de Schonen, 2004). Because the infant face perception literature is relatively devoid of specific reference to infants’ differential categorization and processing of female and male faces, attempts to understand this phenomenon will add to the growing body of literature demonstrating infants’ asymmetrical processing of faces in such areas as species recognition, own-race recognition bias, and categorization of emotion (e.g., Nelson, 1993; Pascalis et al., 2002; Sangrigoli & de Schonen, 2004). Moreover, the results from our prototype studies suggest an alternative explanation than that proposed by Quinn et al. (2002) for the asymmetry in infant processing of female and male faces. They propose that infants may represent both male and female faces as a summary prototype, and have the ability to encode exemplar-specific details of female, but not male, faces. The combined results of our prototype studies, however, suggest that infants do not represent male faces as a summary prototype, and are inconsistent in their ability to encode exemplar-specific details of male faces, as evidenced by their looking to the familiar male faces during the test trials. This issue needs to be examined further. Understanding developmental differences in infant categorization of faces by sex and reasons for this asymmetry is meaningful to social and cognitive development and will contribute to the literature by generating more research to explore this interesting developmental phenomenon.

Asymmetrical learning about male and female faces

The asymmetry in categorization and processing of male vs. female faces may result from the increased difficulty in learning about male compared with female
faces. If infants are indeed developing their knowledge and conceptual understanding of male and female faces at varying rates rather than simultaneously, explanations of underlying mechanisms require not only an understanding of infant face perception, but knowledge of factors affecting category acquisition. Some reasons infants may have trouble forming a category in general are: (1) lack of experience with exemplars from the category (e.g., Homa & Vosburgh, 1976); (2) preferences for other objects (e.g., Eimas & Quinn, 1994); and (3) the range of perceptual differences among the category members (e.g., French, Mareschal, Mermillod, & Quinn, 2004; Oakes & Spalding, 1997; Quinn et al., 1993). Our discussion of each of these factors reflects the assumption that global categorization (i.e., categorization of faces in general) develops prior to basic-level categorization (i.e., categorization of faces into male and female categories) as seen in other forms of categorization (e.g., Mandler & McDonough, 2000; Quinn, 2004; Younger & Fearing, 2000). We posit that this asymmetrical learning is first reflected in infants’ global representation of faces, such that their representation of faces in general is more female-like. As infants then begin parsing faces into the basic-level categories of male and female, they should have a more well-defined category for female than male faces. In the following sections, we will explore each of these factors individually and then discuss how experience, preferences, and perceptual differences among category exemplars might work together to affect infant acquisition of knowledge about male and female faces.

Differential experience with faces

Without experience with particular types of faces, infants cannot form a category for those types of faces (Homa & Vosburgh, 1976). At the other extreme, infants do not need experience with all the faces from a particular category to form the category. Although it is currently unclear how much and what type of experience is necessary for natural category formation to occur, it is apparent that infants have different levels of experience with objects—these differences in experience cause disparities in their formation of categories (e.g., Hayne, Rovee-Collier, & Perris, 1987; Oakes & Spalding, 1997; Shields & Rovee-Collier, 1992).

Infants’ experience with faces seems particularly important for face processing and categorization. Infants seem predisposed to learn about and process a number of different types of faces, but with experience, their categorization and subsequent processing becomes more specialized (e.g., Ludemann & Nelson, 1988; Nelson, 1993; Pascalis et al., 2002; Sangrigoli & de Schonen, 2004). For example, whereas 6-month-olds recognize both human and monkey faces as evidenced by their visual looking after familiarization, 9-month-olds recognize only the human faces (Pascalis et al., 2002). Interestingly, however, 9-month-olds recognize familiar monkey faces shown in familiar and novel orientations as evidenced by their brain activity (event-related potentials), but adults recognize familiar monkey faces shown in the familiar orientation only (Nelson, 1993). Although there are some discrepancies in the two studies with regard to when recognition of other species begins to decrease (likely due to differences in the method and measures used), both studies suggest face processing gets more specialized because of experience with certain types of faces (i.e., experience
primarily with human faces). This phenomenon is not limited to the species of the face; it also relates to infant learning about race and emotion. Infants show better recognition for faces of their own race by 3 months of age (Sangrigoli & de Schonen, 2004). In addition, 7-month-olds’ greater experience with happy than fearful expressions causes them to more easily process and form a category for happy than fearful expressions (Ludemann & Nelson, 1988; Nelson, 1993; Nelson & Dolgin, 1985).

If experience with faces interacts with the brain’s subsequent processing of those faces, then a discrepancy in infants’ experience with female relative to male faces could facilitate easier processing and categorization of female than male faces. Therefore, greater experience with female than male faces should cause infants’ representation of faces in general to be more female-like and should make it easier for infants to learn about the category of female faces when they begin parsing faces into sub-categories. What is the evidence that infants’ experience with male and female faces differs?

Studies investigating the amount and type of interaction between fathers and mothers and their infants support the idea that infants, on average, have more experience with their mother’s than their father’s face. Although fathers’ involvement with their infants has increased in recent years, mothers still tend to be the primary caregiver and spend more time than fathers caring for their infants (Bailey, 1994; Hossain & Roopnarine, 1994; Roopnarine, Brown, Snell-White, & Riegraf, 1995; Russell & Radojevic, 1992; Wille, 1995). Indeed, about 50% of fathers are disengaged from their child or are more likely to serve the role of disciplinarian than caregiver or playmate (Jain, Belsky, & Crnic, 1996). Even in families where the father eventually becomes the primary caregiver, the mother tends to serve as the primary caregiver during the first 6–18 months of the child’s life (Radin, 1988). Although fathers’ involvement with infant care increases as their infants become older (Anderson, 1996; Bailey, 1994; Rustia & Abbott, 1993), it appears that, particularly during early infancy, most babies spend much more time with their mother than their father. Because this discrepancy occurs in caretaking, fathers generally have less physical contact with their infants than mothers, causing a disparity in the type of experience infants have with mother’s and father’s face (Beail, 1985; Ninio & Rinott, 1988; Nugent, Yogman, Lester, & Hoffman, 1988). Given infants’ limited vision prior to 6 months of age, this experience is particularly important because infants can clearly see only individuals at a close proximity to them (Courage & Adams, 1990).

Mothers and fathers also differ qualitatively in their infant interactions. Mothers play more visual games and display more positive affect, whereas fathers engage more often in limb movement games and physical play (Beail, 1985; Forbes, Cohn, Allen, & Lewinsohn, 2004; Nugent et al., 1988; Yogman, 1981). Maternal interactions may direct infant’s attention toward mother’s face, whereas paternal interactions may direct infant’s attention away from father’s face and toward other objects, or their own body. Another difference between mothers and fathers is that mothers tend to be more verbal with infants than fathers (Golinkoff & Ames, 1979; Harrison & Magill-Evans, 1996; Landerholm & Scriven, 1981; Leaper, Anderson, & Sanders, 1998; Yogman, Cooley, & Kindlon, 1988). Because infants apportion more visual attention toward objects when they are accompanied by sound (Balaban & Waxman, 1997), infants may pay greater attention to their mother’s than their father’s face.
because of this disparity in verbal interaction. Therefore, in addition to the differential amount of time that mothers and fathers spend with their infants, these qualitative sex differences in adult interactions with infants could also contribute to infants’ differential experience with mothers’ and fathers’ faces.

Differential experience with their mother and father results in infants’ greater knowledge of their mother than father. At 13 and 20 months of age, infants were more successful at tasks involving mom than dad when they included: visual recognition; spatial location (pointing at mommy or daddy); verbal labeling; featural possession (whose shoe is that?); and identification of sex (Pipp, Easterbrook, & Brown, 1993). Thus, the quantitative and qualitative differences in mother and father interactions may cause infants to acquire more complex knowledge about their mother’s than their father’s face.

Despite the evidence suggesting that most infants have more experience with their mother’s face than their father’s face, it is unclear if this differential experience applies to other female and male faces. Although the literature is lacking on comparing the quantity and quality of infant interactions with non-parental females and males, there is evidence to suggest that infants might have more experience with female than male faces, in general. Many infants receive care outside of their family and the majority of these caretakers are women (Culkin, 1999; Ranck, 1999). In addition, if males do not assist a great deal in the care of their own infant, it is unlikely that they will assist in the care of someone else’s infant. It also seems reasonable to assume that the qualitative differences in how males and females interact with their own infants generalize to the way they act with other infants. These differences between females and males should act to increase infant exposure to female faces and decrease infant exposure to male faces.

Although infants seem to have less exposure to male than to female faces, they still have some experience with male faces, particularly by 6 months of age when their vision has improved and they can more easily see individuals who are not their caregivers and are farther away from them (Courage & Adams, 1990). Differential experience with males and females may not be the only cause of differential categorization and processing of female and male faces. In the series of male prototype studies, Langlois et al. (1997) manipulated the role of experience both within the laboratory (by increasing the number of habituation trials) and outside the laboratory (by testing older infants), but neither of these manipulations affected the results. The ineffectiveness of the experiential manipulations to increase familiarity with male faces suggests that experience alone may not solely account for the asymmetry in infants’ categorization of male and female faces. Some other factor(s) might also contribute to this differential acquisition of knowledge about male and female faces.

Preferences for select faces

In addition to differential experience, infant preferences for female or male faces may affect acquisition of these categories. Infant preferences for some faces over others will affect whether or not those faces elicit the infant’s attention and how long they keep the infant’s attention. If infants prefer to look more at female than male
faces when both are present, it may facilitate finer processing of details in female faces and subsequently enable better within-category discrimination of female faces (Quinn et al., 2002). Such a preference should maintain a more female-like representation of faces in infants and facilitate learning about and categorizing of female faces prior to male faces when infants begin parsing faces into separate sex categories.

By the time infants are 3 months of age, most face perception researchers agree that infant interest in and processing of faces is driven by their experience with faces (e.g., Dannemiller & Stephens, 1988; Kleiner, 1993; LeGrand, Mondloch, Maurer, & Brent, 2001; Morton & Johnson, 1991; Nelson, 2001). Considering that most infants aged 3 months and younger have more experience with their mother’s than their father’s face, and possibly more experience with other female faces, infants should have a more female-like representation of faces. Because infants look longer at faces close to the prototype (Rubenstein et al., 1999), infants should visually prefer female-like to non-female-like faces. Also, the greater amount of vocalizations and physical contact by mothers (and presumably other females) than fathers (and presumably other males) when interacting with infants (e.g., Beail, 1985; Golinkoff & Ames, 1979; Harrison & Magill-Evans, 1996; Ninio & Rinott, 1988; Nugent et al., 1988) could not only affect infants’ experience with faces, but also cause them to prefer female to male faces.

There are several studies in the literature that support the hypothesis that infants prefer looking at female than male faces when paired together. In two studies, 5-, 7-, 9-, and 12-month-olds and 3- to 4-month-olds, respectively, showed stronger recovery to a female than a male test face following familiarization to male faces than they did to a male test face following familiarization to female faces (Leinbach & Fagot, 1993; Quinn et al., 2002). Thus, even after seeing several female faces, infants show no differences in visual interest to familiar (female) and novel (male) category exemplars. Because these studies were conducted independently and used different stimulus faces, these results suggest generalizability of the findings. The asymmetry seen in the test trials of these familiarization/habituation studies led to a direct test of infant preferences for female relative to male faces (Quinn et al., 2002). Three- to 4-month-olds saw female and male faces paired together in a visual preference paradigm. Infants whose primary caregiver was their mother looked longer at female than male faces when paired together. This preference for female faces has been replicated in additional studies with 6- to 7- and 9- to 10-month-olds (Quinn, 2002). In contrast, a small sample of infants with father as their primary caregiver preferred male faces, although it was significant at the one-tailed level only (Quinn et al., 2002). Thus, caregiving appears to be an important factor in driving infants’ visual preferences perhaps because the caregiver’s face is most commonly seen by the infant and is therefore most heavily weighted in their representation of faces.

An alternative explanation is that the skin color of females is lighter than that of males (Zebrowitz, 1997), which could enhance the contrast between the eyes and mouth and the rest of the face for females. This greater contrast could increase infant interest in female faces because their facial features would be easier for the infant to
see (Banks & Salapatek, 1981). Quinn et al. (2002) investigated this possibility by showing 3- to 4-month-olds inverted pictures of females with inverted pictures of males because a preference due to greater contrast should be present regardless of whether the pictures are presented upright or inverted. They found, however, that infants demonstrated preferences for female relative to male faces when the pictures were presented upright only, not when the pictures were inverted. Thus, differences in contrast between female and male faces seem an unlikely explanation for differential interest in female and male faces. Caregiving appears to be a more important factor for driving interest in faces and this interest can affect face categorization.

The visual preference for female relative to male faces when infants see the two types of faces paired together may seem contradictory to our meta-analysis results showing that infants look longer at male than female faces when infants see only male or only female faces. Because the results from these studies are limited to a laboratory setting, it is unclear how these looking times translate within infants’ natural environments. The results suggest, however, that infant interest in female faces will predominate in settings with both female and male faces. Whereas visual preferences for female faces when both females and males are present may permit finer and more expert processing of female facial details (Quinn et al., 2002), longer looking at male faces when only males are present may result from their novelty and the fact that more time is needed to process them. Given infants’ more limited experience with male than female faces, the tendency to spend a greater amount of time viewing male faces when only males are present suggests one mechanism aiding infant learning of males.

Range of perceptual differences among category exemplars

Evidence from the infant categorization literature demonstrates the importance of considering how the range of perceptual differences among category exemplars affects categorization (e.g., Bauer, Dow, & Hertsgaard, 1995; French et al., 2004; Oakes, Coppage, & Dingel, 1997; Oakes & Spalding, 1997; Quinn et al., 1993). If there is high variability among category exemplars, infants are more likely to experience an atypical than a prototypical category member. In contrast, if there is low variability among the range of exemplars, infants are more likely to experience a prototypical than an atypical category member. The greater the perceptual similarity or prototypicality among category exemplars, the easier it is for infants to form a category because it is easier to form and differentiate between categories when the within-category similarity is maximized and the between-category similarity is minimized (e.g., Bauer et al., 1995; Oakes et al., 1997; Oakes & Spalding, 1997; Quinn et al., 1993; Roberts & Horowitz, 1986). Evidence from the sexual selection, face perception, and anthropology literatures suggests that male faces are more variable and thus less similar than female faces (e.g., Alley & Hildebrandt, 1988; Darwin, 1871/1981; Farkas, 1981; Farkas & Munro, 1987; Johnston, Kanazawa, Kato, & Oda, 1997; Symons, 1979, 1987; Vokey & Read, 1988, 1992). This range of the female and male face categories is therefore important to consider in understanding infants’ asymmetrical learning of female and male faces.
Sexual selection theory posits that females evolved to be more attractive than males and that males evolved to vary more than females (Darwin, 1871/1981; Symons, 1979, 1987). Specifically, the theory posits that females have been sexually selected for beauty by their mates, and that over time this selection pressure has caused women to become more beautiful than men. This sex difference in the selective advantage of physical attractiveness for females could have resulted in less variability in female than male faces (Alley & Hildebrandt, 1988). Because of the greater emphasis on female choice in selecting a mate, Darwin proposed that a number of variations in males may arise through attempts to attract females. Indeed, he found that numerous measurements of males and females from various races demonstrated that males differed more from each other than females differed from one another.

There may also be more variability in male than female faces because of secondary sex characteristic differences. One theory is that male facial structure is more affected by sex hormones (e.g., testosterone) than female facial structure (Thornhill & Gangestad, 1993, 1999). Because the level of testosterone can vary greatly among individual males (e.g., Salvador, Suay, Martinez-Sanchis, Simon, & Brain, 1999; Silverman, Kastuk, Choi, & Phillips, 1999), there should also be high variability in how male facial structure is affected. Although empirical research is needed to support the idea that testosterone affects male facial growth, something about secondary sex characteristics differentiates male from female faces because these differences are not apparent in pre-pubescent boys and girls (Enlow, 1990).

Empirical evidence supports these theories. Adults rate a greater percentage of female than male faces as typical (e.g., Vokey & Read, 1988, 1992). The greater typicality of female faces suggests that they deviate less from the average of the population than male faces. Adults also rate female faces as more attractive than male faces overall (e.g., Hume & Montgomerie, 2001; Oliver-Rodriguez, Guan, & Johnston, 1999; Vokey & Read, 1992). Attractiveness is positively correlated with typicality (O'Toole et al., 1994; Vokey & Read, 1992), and this relationship is in accordance with Langlois and Roggman's (1990) finding that faces whose configurations are closer to the mean of the population are more attractive. The higher ratings of typicality and attractiveness for females compared to males support the idea that female faces are less variable and more prototypical than male faces.

Further support for the idea that male faces are less prototypical and more variable than female faces comes from a study assessing adults’ ability to classify whether or not male and female adult and child faces and “scrambled” versions of the same faces were indeed a face (Johnston et al., 1997). Participants required more time to classify a male face as a face than to classify a female face as a face, suggesting the greater typicality of female than male faces. These findings were replicated using both Japanese and Caucasian faces, suggesting typicality of female faces and variability of male faces may be universal.

Finally, our own analysis of female and male facial measurements demonstrated that male faces are indeed more variable than female faces. Using 18-year-olds’ basic facial measurements and standard deviations as reported in Farkas (1981) and 18- and 26-year-olds’ facial measurements and standard deviations as reported in Farkas and Munro (1987), we performed 1-tailed, paired-samples t tests to determine
whether the standard deviations of male facial measurements were significantly larger than the standard deviations of female facial measurements. These researchers collected facial measurements from these age groups and younger individuals to provide general population baseline data for craniofacial development (Farkas, 1981) and to assess age, ethnic, and sex differences in facial proportions (Farkas & Munro, 1987). For the Farkas (1981) data, the results demonstrated significantly more variability of the measurements of males than females: the mean of the standard deviations of the facial measurements was 4.6 for males and 4.4 for females, \( t(130) = 2.67, p = .005 \). For the Farkas and Munro (1987) data, we found significantly more variability in facial and head measurements of males than females: the mean of the standard deviations of the facial measurements was 3.8 for males and 3.5 for females, \( t(44) = -2.18, p = .02 \); and the mean of the standard deviations of the head measurements was 4.4 for males and 3.8 for females, \( t(28) = -2.10, p = .02 \). There were, however, no significant differences in variability of individual features, such as the lip and mouth region, nose, ear, or eyes. In general, the results suggest that when there are variability differences, it is in the direction of male faces being more variable than female faces.

Despite much evidence supporting the idea that male faces are less prototypical than female faces, there is counter evidence suggesting that female faces may be less prototypical than male faces. Studies that assessed male and female face recognition using neural network modeling found that male faces were easier for the neural network model to classify and were closer to the general face average than female faces, thus suggesting greater variability of female faces (Abdi, Valentin, Edelman, & O'Toole, 1995; Abdi, Valentin, & O'Toole, 1997; O'Toole et al., 1998). A follow-up study, however, showed that the sex of face contributed very little to the variance in face recognition (less than 1%), and that the model was generally not classifying faces in the same way as human adults (Edelman, Valentin, & Abdi, 1998). Thus, it is unclear how appropriate it is to generalize the neural network model results to human face perception.

Theory and empirical research support the idea that male faces are less prototypical and more variable than female faces (Alley & Hildebrandt, 1988; Darwin, 1871/1981; Hume & Montgomerie, 2001; Johnston et al., 1997; Oliver-Rodriguez et al., 1999; Symons, 1979, 1987; Thornhill & Gangestad, 1993, 1999; Vokey & Read, 1988, 1992). Greater variability in male than female faces is consistent with other research finding greater variability in males than females in intelligence and cognitive, perceptual, and reasoning abilities (Feingold, 1992). This variability should make it more difficult for infants to form male than female face categories.

Although the notion that less variability and greater prototypicality of category exemplars aids categorization has received much support from the literature with infants, older children, and adults (Bauer et al., 1995; Ellis & Nelson, 1999; Medin & Schaffer, 1978; Oakes et al., 1997; Oakes & Spalding, 1997; Quinn et al., 1993; Roberts & Horowitz, 1986; Rosch et al., 1976), recent evidence provides an alternative explanation. French et al. (2004) found that category inclusiveness, rather than category variability, may explain the findings from one study that found 3- to 4-month-olds had difficulty forming a category for dogs, but not cats, because there was
greater variability among the dogs than the cats used in this study (Quinn et al., 1993). Through neural network modeling and subsequent testing of infants, French et al. (2004) demonstrated that featural inclusiveness was more important in understanding these results than variability of the category. More specifically, they measured several features of the dogs that infants saw in Quinn et al.’s (1993) study (e.g., head length and width) and compared these measurements to the same features of the cats. They found that most of the cat measurements were inclusive within the dog measurements, but did not find the reverse. This inclusiveness within categories may explain the asymmetrical categorization of cats and dogs.

Although featural inclusiveness may explain Quinn et al.’s, 1993 findings better than the variability of the category, it is unclear if this explanation generalizes to other findings in the literature suggesting the importance of prototypicality for infant categorization (e.g., Bauer et al., 1995; Oakes et al., 1997; Oakes & Spalding, 1997; Roberts & Horowitz, 1986). It is also unclear if inclusion is mutually exclusive from the notion of variability. If one category is more variable than another category, it does not necessarily need to include a range of features that overlaps with the less variable category. If, however, one category’s features are included within the range of features of another category, by definition the former category is going to be less variable than the latter category. Thus, inclusiveness seems to be one form of variability rather than an orthogonally distinct explanation. Given that male facial measurements are generally larger than female facial measurements (Enlow, 1990; Farkas, 1981; Farkas & Munro, 1987), female facial measurements are more likely to be included within male facial measurements than vice versa. This inclusiveness, therefore, may be one component of the greater variability of male than female faces.

Combined effects of experience, preferences, and variability

Differential experience, preferences, and variability may all individually affect infant categorization of male and female faces. The combined effects of each of these factors are likely to contribute to: (1) infant formation of a category for faces in general based primarily on female faces and represented by a female-like configuration; and (2) significant developmental differences between when infants form well-defined categories for female and male faces. Early experience with females may cause infants to prefer female faces if infant interest is driven by such experience (e.g., Dannemiller & Stephens, 1988; Kleiner, 1993; Morton & Johnson, 1991; Nelson, 1993, 2001; Pascalis et al., 2002; Quinn et al., 2002). In turn, preferences for selected faces will increase infants’ experience with those faces. Thus, the effects reinforce each other.

If infant preferences for and greater experience with female than male faces occurs early in development, these effects will augment the effects of the greater variability of male faces. Although category variability makes it difficult to learn that category, it becomes even more difficult when exemplars are encountered infrequently (Homa & Vosburgh, 1976). Thus, the less experience infants have with males because of more interaction with and greater preferences for females, the more variability will affect infant formation of a category for female and male faces.
Proposed developmental trajectory of face category acquisition

Based on the literature we reviewed and the reasons we proposed for the asymmetry in infant categorization of faces, we now outline a developmental trajectory describing how infants form a category for faces in general as well as sub-categories related to sex of face. We discuss this trajectory with a few caveats. First, much work is needed in this field to better understand how infants are learning about and categorizing faces, so there are gaps in our proposed trajectory that need to be filled and suppositions that need to be supported or refuted. Second, the proposed trajectory is based on the assumption that the infant’s primary caregiver is a female, at least during the first year of life, because these infants represent the majority of the population and thus the majority of the participants from studies conducted on infant face perception. For infants with male primary caregivers, it is currently unclear whether a similar asymmetry exists and how the trajectory may differ. Last, sex categorization abilities are not an all or none phenomenon (we thank one of our reviewers for suggesting this aspect of categorization). Throughout development, everyone experiences faces that do not fit well into their existing categories (e.g., androgynous-looking individuals, transgendered individuals). As a result, category boundaries should adjust according to experience. Because of limited experience with atypical exemplars of sex categories, however, even adults are not 100% accurate in categorizing an individual’s sex of face (e.g., Bruce et al., 1993; O’Toole et al., 1998).

During the first month of life, infants begin to learn about faces while spending most of their waking time with their mother. Thus, they first learn to recognize and discriminate her face from other female faces (e.g., Field et al., 1984; Pascalis et al., 1995). By the third month, these discrimination abilities extend to other female faces (Barrera & Maurer, 1981a), at least to female faces of the same race (Sangrigoli & de Schonen, 2004). As a result of experience with various faces, infants should begin to form a representation of faces and a rudimentary category for faces (de Haan et al., 2001). This representation should be most heavily weighted with the mother’s face and therefore specific to the human species, most representative of the mother’s race, and primarily female-like, so that it guides infants’ attention toward other female-like faces (Pascalis et al., 2002; Quinn, 2002; Quinn et al., 2002; Sangrigoli & de Schonen, 2004). As infants experience new faces, their representation should adjust accordingly. If, however, infants continue to more frequently experience and prefer looking at female than male faces, their representation of faces will remain female-like. Particularly during the first 6 months while the infants’ vision is still relatively limited (Courage & Adams, 1990), the people engaging in the caregiving of the infant should most influence the face representation.

As infants develop face representation, it should not only be female-like, but also attractive because it is a composite of the faces they experience (de Haan et al., 2001; Langlois & Roggman, 1990; Rubenstein et al., 1999). Infants may consequently begin to group faces that are similar to this representation (e.g., attractive female faces and perhaps attractive, feminine-looking male faces). This statement does not mean that infants do not recognize other faces as a face (Younger & Fearing, 1999), but that the first sub-category they develop is for faces that are “face-like” (i.e., faces that most
closely match their representation of faces). Such a hypothesis would explain why infants between the ages of 6–11 months are good at categorizing only very stereotypical female (very female-like) and very stereotypical male (very non-female-like) faces in habituation studies, but have difficulty categorizing less stereotypical-looking faces—faces that have “fuzzier” boundaries between them than the extremes of very female-like and very non-female-like faces (Leinbach & Fagot, 1993; Newell & Strauss, 2002; Newell et al., 2003). The hypothesis also provides an explanation for why 6-month-olds demonstrate the ability to categorize female faces as high attractive (very female-like) and low attractive (less female-like) within a familiarization paradigm (Ramsey, Langlois, Hoss, Rubenstein, & Griffin, 2004). Furthermore, this hypothesis helps explain why infants can form a prototype for female faces, but not male faces and why they prefer high attractive (prototypic) female faces to low attractive (non-prototypic) female faces (e.g., de Haan et al., 2001; Langlois et al., 1987, 1991, 1997; Rubenstein et al., 1999). Within a laboratory setting, it should be much easier for infants to form a prototype that is similar to a pre-existing representation of faces from their natural environment than one that is quite different from this pre-existing prototype.

Because very little infant research has used “non-female-like” faces as stimuli, it is much less clear how infants are categorizing/sub-categorizing these faces, but the greater variability of male than female faces provides some potential hypotheses. Greater variability of male than female faces suggests that infants should more often experience atypical male face exemplars than atypical female face exemplars, making it difficult to form category boundaries for male faces. As a result, infants may form boundaries for the category of male faces that: (1) are too broad and include some members of the contrasting category; or (2) are too narrow and contain only a subset of the members of that category (Oakes & Spalding, 1997).

If infants develop a category for male faces that is too broad, the neighboring members of the contrasting category that infants are likely to include (incorrectly) are low attractive (less female-like) female faces. Because attractiveness is related to typicality (O’Toole et al., 1994; Vokey & Read, 1992) and averageness (Langlois & Roggman, 1990; Langlois, Roggman, & Musselman, 1994), unattractive female faces will be the least representative of the category of female faces and therefore are most likely to be miscategorized into a broad category of male faces. Thus, this hypothesis suggests that infants are categorizing non-female-like faces as one sub-category. Miscategorizing unattractive female faces into a “male” category is a testable hypothesis and should be examined in future research.

On the other hand, the variability and differences due to secondary sex characteristics in male faces may cause infants to form subsets of the category of male faces (e.g., subsets based on masculinity level or presence/absence of facial hair). Subset categories for males may persist because as infants compare the exemplars they categorize, they tend to disregard non-category members even when they are similar-looking to the category members (Younger, 1993). Therefore, it may take some time before infants group all male faces together (i.e., group male faces by their sex rather than facial hair or masculinity).
This broadening or narrowing of the male face category can hinder infants’ ability to correctly develop categorical differentiation of male faces and the two phenomena may even work in conjunction with one another. Infants may first form categories for female-like and non-female-like faces. As infants begin to differentiate the non-female-like face category into sub-categories, they may parse out the atypical female faces from the male faces, as well as parse the male faces into various sub-categories based on differences in secondary sex characteristics.

Ill-defined category formation of male faces also has implications for the prototype(s) infants will form. The prototype should represent the average of the infants’ category, but if the category is too broad and includes faces other than males, or is too narrow and is divided into subsets of male faces, infants will form prototypes for male faces that represent these broad or subset categories. These prototypes will differ from the adult version of male face categories.

If infants’ male face category is too broad or too narrow, when does the adult-like version appear? The variability in male faces is a perceptual cue; therefore, infants should begin to appropriately categorize male faces when they begin to rely more on knowledge-based conceptual cues and not solely on perceptual cues when categorizing male faces. Because infants are good at detecting correlations in their environment (e.g., Greco, Hayne, & Rovee-Collier, 1990; Safran, Johnson, Aslin, & Newport, 1999; Younger, 1992), some conceptual cues may be infants’ understanding that certain behaviors or traits tend to be associated with males. For example, one such cue may be the understanding that low frequency male voices are associated with male faces—this awareness appears to occur and generalize to several types of male faces around 18 months of age (Poulin-Dubois et al., 1998). Infant labeling of males and females may also signify that their categories have changed from perceptual discrimination of the two sexes to explicit knowledge of the categories (Fagot & Leinbach, 1993). Most children can label the sex of their parents by 16 months of age, but they generally do not succeed at labeling other adults until sometime between 24 and 40 months of age (Fagot & Leinbach, 1993). Learning label categories is generally facilitated by children’s ability to physically sort objects into groups (Gopnik & Meltzoff, 1987) and most children sequentially sort by sex around 22 months of age (Johnston, Bittinger, Smith, & Madole, 2001), so naming of sex categories should indicate that children are developing more accurate sex categories. Finally, toddlers’ knowledge of their own sex is apparent in their own self-labeling between 26 and 36 months of age (Katz & Kofkin, 1997; Weinraub et al., 1984) and prior to this labeling, they demonstrate visual preferences around 18 months of age for toys traditionally associated with their sex (Serbin, Poulin-Dubois, Colburne, Sen, & Eichstedt, 2001). With the onset of gender identity, sex differences in visual interest in and processing of male faces may emerge as male children become more motivated than female children to learn about male faces and other gender-relevant characteristics (Martin, Wood, & Little, 1990). This early sex difference may help to explain why adult females demonstrate an advantage in recognizing other adult female faces (e.g., Lewin & Herlitz, 2002; Shepherd, 1981): Although both female and male infants will show greater
expertise in processing female faces during early infancy, males should adjust their visual preferences more toward male faces as they form a gender identity, whereas females should maintain their visual preferences for female faces and thus strengthen their expertise in processing female faces.

This conceptual knowledge should result from infants’ experience with males, which we have already noted is less frequent than their experience with females. Thus, infants may first adjust their category for female-like faces by beginning to group high feminine, high attractive female faces with less feminine, less attractive female faces (and all those faces in between these extremes) into a category for female faces before they adjust their category for non-female-like faces into a category for male faces. Note we are not speculating that the female face category replaces the female-like face category nor are we speculating that the male face category replaces the non-female-like face category; rather, the two categories may co-exist with the female and male face categories providing the basis for gender stereotypes and the female-like and non-female-like categories providing the basis for femininity/masculinity stereotypes.

Conclusions

Learning to categorize a face as female or male is an important part of development for both social and evolutionary reasons. Results from empirical research suggest that infants categorize and process female and male faces differently in that they demonstrate an advantage in processing female faces (e.g., Poulin-Dubois et al., 1994, 1998; Quinn et al., 2002; Younger & Fearing, 1999; and compare de Haan et al., 2001; Rubenstein et al., 1999 to Langlois et al., 1997). Differential experience with, preferences for, and differences in variability of female and male faces, on average, may all contribute to this asymmetry. We have proposed a developmental trajectory for categorization and sub-categorization of faces that suggests infants have an early representation of faces that is female-like and that they demonstrate a developmental lag in the categorization of male faces. Although these hypotheses clearly need to be tested, the ideas can be empirically supported or refuted. Because of this asymmetry and concerns that young infants may respond best when the stimulus faces are sex-stereotypical-looking females, researchers studying infant face perception should include both male and female faces that range in stereotypicality in their studies. At a minimum, researchers using only sex-stereotypical-looking female faces as stimuli should be careful not to generalize their results to all faces. Until studies using only stereotypical-looking female faces as stimuli are replicated using male faces and non-stereotypical-looking female faces, the literature is ill-equipped to reveal a single, overall developmental ability. Future studies using a range of both male and female faces are important for understanding the development of face recognition, emotion recognition, gender and other appearance-based stereotypes, and infant use of facial cues during social interactions.
References


