GENDER DIFFERENCES IN FACE RECOGNITION: THE ROLE OF INTEREST AND FRIENDSHIP

Johanna Lovén

Women outperform men in face recognition and are especially good at recognizing other females’ faces. This may be caused by a larger female interest in faces. The aims of this study were to investigate if women were more interested in female faces and if depth of friendship was related to face recognition. Forty-one women and 16 men completed two face recognition tasks: one in which the faces shown earlier had been presented one at a time, and one where they had been shown two and two. The Network of Relationships Inventory was used to assess depth of friendships. As hypothesized, but not statistically significant, women tended to recognize more female faces when faces were presented two and two. No relationships were found between depth of friendships and face recognition. The results gave some support for the previously untested hypothesis that interest has importance in women’s recognition of female faces.

Imagine not being able to recognize your friends’ faces, and having to rely on other cues to know who you were looking at. This notion ought to illustrate that faces are of critical importance in our everyday lives, and so is our recognition memory for them. Have you ever considered in which areas you have expertise in identification and recognition? Faces are surely one area of expertise for most of us, as well as other objects we are interested in, like birds or cars. Interest could be an important factor in face recognition: it may make some people better at it, as more attention may be given to the stimulus if one is interested in it. The aims of this study were to investigate the role of interest in women’s recognition of female faces, through altering one of the standard face recognition procedures, and to see if there were any relationships between face recognition and reported depth of friendships. Interest for other people in general could be related to face recognition performance, and this might be reflected in persons’ friendships. It could be speculated that people who are generally more interested in other people report a higher grade of closeness in friendships, and that this might be related to a better memory for faces.

Gender differences in cognitive abilities are, as many researchers have noted (see for example Halpern, 2000), a hot topic, and especially the proposed reasons for the findings. The magnitude of effect sizes of gender differences varies depending on which ability that is measured. Women have been found to outperform men on some tests of memory and verbal abilities (Stumpf, 1995), and advantages favoring men have been found on most visuospatial tasks (Voyer, Voyer, & Bryden, 1995). In addition, the direction of gender differences can vary between women and men even within a given ability (i.e. verbal abilities and episodic memory), and also across age. Women have been found to
outperform men on verbal episodic memory tasks and men have been found to perform on a higher level than women on visuospatial episodic memory tasks (Lewin, Wolgers, & Herlitz, 2001).

Effect size is a useful measure to assess the magnitudes of these differences. Cohen's $d$ can be used when comparing group means and Pearson's $r$ can be used for magnitude and direction of effect (McCartney & Rosenthal, 2000). The $d$-value is the difference between two group averages, divided by the combined standard deviance and is therefore a standardized difference between the groups' means (McCartney & Rosenthal, 2000). There are different interpretations of the importance of a computed $d$-value. McCartney and Rosenthal (2000) defined $d = .20$ as small, $d = .50$ medium, $d = .80$ large (in accordance to Cohen's definition). An $r$ sized .10 was regarded as small, .30 medium, and .50 large (McCartney & Rosenthal, 2000). A small $d$-value does not mean that the size of the difference does not have practical importance. Consider for example a very small $d$ value, which may be of a huge importance if people’s lives can be saved through using a certain method.

Verbal abilities are described as the ability of using language (Halpern, 2000), and can be tested for example with tests of reading and writing abilities, fluency (generation of words), speech articulation, and grammar knowledge (Halpern, 2000; Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003). Women tend to perform better than men on most verbal tasks (Halpern, 2000), for example on letter fluency (e.g., Maitland, Herlitz, Nyberg, Bäckman, & Nilsson, 2004). The effect size of the gender differences in verbal abilities varies, from quite small or non-existing $d$ sizes of .02 for vocabulary to more moderate sizes of .33 for speech production (e.g., Hyde & Linn, 1988).

Visuospatial tasks aim to test the ability to mentally rotate figures or to distinguish relationships between objects and shapes, and processing of spatial, nonverbal, information (Halpern, 2000). On one type of visuospatial tasks, the mental rotation test, where rotated block figures are compared to a standard block figure to determine which figures are the same as the standard one, men tend to outperform women (Voyer et al., 1995). This is the test where the largest and most stable gender difference can be found among the visuospatial abilities, with moderately to large effect size, for example, a $d = .56$ was found in a meta-analysis by Voyer et al. (1995).

Memory and Gender Differences

Memory is not a unitary construct (Halpern, 2000), since it can be classified into different subsystems: short-term memory and long-term memory. Short-term memory has also been called working memory, a limited information storage, which briefly can store and process information (Baddeley, 2003). Long-term memory is divided into declarative and non-declarative memory subsystems. Declarative memory concerns factual knowledge (Passer & Smith, 2004). It can be split in two subcategories: episodic memory, which has been defined as the memory storage of information about our personal experiences and episodes (Passer & Smith, 2004), and semantic memory, which is involved with general knowledge and vocabulary (Maitland et al., 2004). Semantic memory is often tested with general knowledge tests, for example vocabulary or synonyms tests, and typically, no
gender differences in general knowledge and vocabulary are found (Maitland et al., 2004). Tests that can be administered to measure verbal episodic memory are for example recall of actions and/or nouns. Women tend to perform on a higher level than men on verbal episodic memory tasks (Lewin et al., 2001; Maitland et al., 2004). To test nonverbal episodic memory, face recognition tests can be used (Lewin et al., 2001).

**Face Recognition**

Faces might be seen as any other object and it could be reasonable to expect that faces are processed in the visual system in the same way as common objects (Farah, 2000). This idea is a bit problematic, as there are plenty of evidence of a specialized recognition system of faces, or at least differences in depth of familiarity and expertise between recognition of faces and common objects (Farah, 2000).

From both behavioral and neuropsychological data there is evidence for and against the theory of a separate face and object recognition systems have been generated. To begin with, the behavioral data that support the idea of a separate system includes findings of infants’ perception of faces. Newborn infants have been found to treat faces differently from non-face objects, for example to follow a moving face longer than other objects (Johnson, 1997, in Farah, 2000). On the other hand, this biologically grounded predisposition seen in infants may just reflect that faces could be the first kind of stimuli the human recognition system is set to process (Farah, 2000). Studies of pattern perception, that is, inverted faces and objects in relation to their normal positions, have found that inverted faces are harder to recognize for adults than upside-down non-face objects (Valentine, 1988, in Farah, 2000). Bindemann, Burton, Hooge, Jenkins, and De Haan (2005) found that faces were harder to shift attention from, if the faces were shown in upright position. Faces shown upside down did not have the same response delaying effect on the task the participants were to perform, and seemed to have the same effect on response time as non-face objects. These data could support the hypothesis of a specialized face processing system (Bindemann et al., 2005). However, the authors noted that an attention bias does not have to mean that faces are processed differently from other objects. It just seems as if humans tend to have an attention bias directed to faces, something which has been found for alcoholics and cigarette smokers if shown relevant stimuli (see review in Bindemann et al., 2005). The reasons for an attention bias for faces could be to quickly be able to notice changes in other people’s faces as they hold socially important information, or to aid the processing of faces when one makes judgments of face familiarity (Bindemann et al., 2005).

Evidence of a specialized function for recognition of faces in human brains has been studied in primates. Baylis, Rolls, and Leonard (1985) studied brain activity in macaque monkeys and the results indicated that certain neurons in the temporal lobe were activated mainly when faces of other monkeys were shown. Another indication of a specialized area for recognition of faces are individuals with prosopagnosia, a deficit where the person has problems in recognizing faces (Anderson, 2005).

In addition, brain-imaging studies with fMRI-data have shown that a certain area in the temporal lobe, called the fusiform (face area) gyrus, is activated when faces are shown
McCarthy et al. (1997) noted that the processing of faces seemed to differ from the processing of non-face objects in the visual system. The authors compared how the brain responded to face stimuli in relation to non-face objects or inverted faces. The conclusion drawn by the authors was that the perception of faces to a degree is handled differently than the perception of other objects, since faces activated different areas than non-face objects did. One reason why responses to inverted faces compared with responses to upright faces is important is because one key to understanding face processing may lie there. If only upright faces are processed specially, in relation to other non-face objects and inverted faces, then this supports the idea of a specialized system for face recognition: faces shown upside down seem to be processed like non-face objects.

Other results supporting this hypothesis have been found in some studies where recognition of inverted faces has been tested (Anderson, 2005). A decrease has been found in the fMRI response in the fusiform gyrus when faces are shown upside down (Haxby, Ungerleider, Clark, Schouten, Hoffman, & Martin, 1999, in Anderson, 2005). Anderson (2005) described one original study by Yin (1969, in Anderson, 2005), where it was found that recognition of faces was better than recognition of other objects, for example houses, when these faces and objects were presented in an upright position. A reduction in recognition was found when faces were presented upside down, which was not the case for non-face objects. Based on this finding, it seems as humans are very accustomed to recognizing faces (Anderson, 2005). Further, a predisposition in humans to identify whole faces can be argued, and this may have an evolutionary basis (Anderson, 2005). This has been tested using different measures, for example comparing relative recognition of faces to houses, and to parts of houses and parts of faces (Tanaka & Farah, 1993, in Farah, 2000). Results from the study reviewed by Farah (2000) showed that recognition of facial parts (e.g. a nose) was easier to recognize if presented in a whole faces, indicating that faces are processed more holistically than other objects.

On the other hand, there are results which do not support the hypothesis of specialized face recognition mechanisms, and an alternative explanation has been proposed: face recognition is rather one kind of making expertise judgments (Anderson, 2005). The needed level of categorization differs between objects in general and for faces, because faces need more finely grained distinction and thus subordinate-level classification to recognize separate individuals’ faces (Farah, 2000). There are results which point to the function of the fusiform gyrus to be involved in fine-grained judgments and recognition of very familiar objects (Farah, 2000). Neuropsychological data from experts making fine-grained judgments in their area of expertise has been studied in for example bird experts and car experts (Gauthier, Skudlarski, Gore, & Anderson, 2000). The participants were given tasks with photographs of faces, familiar objects such as a glass, a table, and expertise targets of different cars and bird species. Gauthier et al. (2000) reported that the most remarkable result of their study was the finding of a very strong relationship between the tests of expertise in objects and the responses in the right fusiform gyrus to cars and birds. Whatever the reasons for our face recognition ability are, we are definitely experts in this area.
Face Recognition: The Procedure
Face recognition tasks are generally tested with similar procedures, but they may differ depending on what area that is studied. For identification studies and other related areas of eye-witness research, other methods than the yes-no recognition task may be used, like the sequential line-up procedure, which is similar to the yes-no paradigm, or the simultaneous line-up identification procedure (Sporer, 2001). The yes-no standard procedure includes a number of photographs of faces including men, women or both genders presented one at a time in a sequential manner with for example a PC projector. First the target faces are presented to the participants and then some time passes until the participants are presented with the recognition task. This task is a forced yes or no task, which means that the participants must score yes or no recognition for each face. If any background is shown it is to be the same for all faces included, furthermore, no beards, no exceptional hair-styles, eye-glasses or jewelry are to be shown (see for example Lewin & Herlitz, 2002; Wright & Sladden, 2003). These conditions are used to ensure that no external cues are available.

Gender Differences in Face Recognition
Some findings indicate a general female advantage in the ability to recognize faces (e.g. Herlitz, Nilsson, & Bäckman, 1997; Lewin et al., 2001; Rehnman & Herlitz, 2006). Herlitz et al. (1997) found a $d = .27$ for face recognition in adults, a medium sized effect indicating a general female advantage in face recognition. Similar results have been found for children aged 8-10 years, where girls generally recognized more faces than boys did (Rehnman & Herlitz, 2006).

Lewin and her colleagues (2001) hypothesized that their results indicated a general female advantage in episodic memory, and that women’s advantage may lie in women’s better verbal abilities. That is, if the material can be verbalized, then females may benefit from this, for example on face recognition tasks. A face could, according to Lewin et al. (2001), get a verbal label (young, handsome woman) and thus be more easily remembered. This hypothesis was tested in 2002 by Lewin and Herlitz. Faces were presented in one condition for three seconds and in a second condition for one second. It was thought that the shorter presentation time would not make verbal labeling of the face possible. Lewin and Herlitz (2002) found that women’s performance on face recognition tasks was not influenced by their verbal ability, because the size of the gender difference in favor of women was not affected by the short presentation time.

Findings of a gender difference, favoring women, in face recognition has been tried to be explained by gender-related differences in interest, meaning that women in general would be more interested in other people than men (Goldstein & Chance, 1970, in Lewin & Herlitz, 2002). The findings of Rehnman and Herlitz (2006) were proposed to be caused by gender-related differences in interest. Faces in general could be objects of greater interest to women, and that this might have a biological basis. A study of one-day old infants showed gender differences in time spent looking at objects. Girls were more interested in a face and boys in a mobile (Connellan, Baron-Cohen, Wheelright, Batki, & Ahluwalia, 2000).
Own-bias Effects in Face Recognition

Results from Lewin and Herlitz (2002) and others (Rehnman & Herlitz, 2006) also indicate that above the general gender difference in face recognition women’s advantage in face recognition may be driven by their better recognition of female faces. Women in the ages of 20-40 years were better than men on recognizing female faces, but no gender difference in recognition of male faces was found (Lewin & Herlitz, 2002). This effect, labeled own-gender bias, seems to be well established with women and girls (Lewin & Herlitz, 2002; McKelvie, 1981; Rehnman & Herlitz, 2006; Wright & Sladden, 2003), showing that women tend to be better at recognizing female faces. For men, however, an own-gender bias is not at all established as only few studies have indicated an effect of this kind (Wright & Sladden, 2003).

Own-biases are shown in interaction effects which have been found both for same gender, sex of the viewer in combination with the gender of the target face, and same race, ethnicity of the viewer combined with the ethnicity of the target face (Wright & Sladden, 2003). Own-gender bias is the tendency found for people to have a better memory for same-gender persons (Wright & Sladden, 2003). Meissner and Brigham (2001) described own-race bias as the tendency for people to be better at recognizing persons of one’s own race and ethnicity.

Reasons for an own-race bias have for example been proposed to lie in cognitive processes like interest and attention (Meissner, Brigham, & Butz, 2005). Meissner et al. (2005) reviewed suggestions from the stereotyping literature of a categorization response to other-race faces, which may explain the own-race bias. It might be that a stereotypical encoding process operates in the perception of other-race faces and categorizes the face automatically as not belonging to the same race. This categorization response could lead to cognitive disregard, which means that the attentional resources are unfocused in the encoding process, thereby resulting in a reduced recognition of other-race face. It could also be the case that levels of processing in encoding operate in recognition of faces, and that other-race faces are less deeply processed than own-race faces (Meissner et al., 2005). Simplified, support for the idea that deeper encoding may lead to better memory comes from findings of the perception of own-race faces. Meissner et al. (2005) found that own-race faces tended to be perceived as more familiar and memorable than other-race faces.

Contact with other ethnicities has also been proposed as a reason for the own-race bias, as familiarity could be seen as one explanation for the own-race bias (Meissner et al., 2005). In a meta-analysis of the own-race bias, Meissner and Brigham (2001) found that contact mediated the magnitude of the own-race bias, supporting the findings of less own-race bias in individuals from integrated populations, in comparison to people from less integrated populations.

Other hypotheses are the theory of perceptual learning and the configural-featural hypothesis (Meissner & Brigham, 2001). The hypothesis of perceptual learning, assessing the cognitive mechanisms thought to operate the own-race bias, holds that with practice and experience one learns to direct attention toward distinct cues in a set of objects that
will help the ability to separate them. Further, if one can process faces more accurately and faster, this ought to reflect how much previous information one has of the faces, and how much attention one has dedicated to the right (invariant) details of the stimulus.

Wright and Sladden (2003) found an own-gender bias, with both men and women being more accurate in the recognition of same-gender faces. In the same study, hair was found to be important in the identification of same-gender faces. Shapiro and Penrod (1986) conducted a meta-analysis of facial identification studies. They found an own-gender bias for hit answers (number of correctly identified faces), showing that same-gender targets were more easily identified. No difference was found in false alarms rates (number of faces falsely identified as previously seen).

Wright and Sladden (2003) propose an evolutionary explanation of the own gender bias. The observed results may be due to the evolutionary basis of women and men higher interest in same gender persons due to competition for mating and a need to weigh up their rivals. In many cases, and as it certainly is in this case, evolutionary explanations are highly speculative. One might argue that it would be of more importance, from an evolutionary point of view, for women to evaluate men due to selection of mating partner, and vice versa for men, and therefore have a better memory for men than for other women’s faces. Another explanation for an own-gender bias proposed by Wright and Sladden (2003) is exposure to non-pornographic magazines and photographs, as these are often of the same gender as the target readers. This explanation is mentioned as a possible reason for an own-gender bias seen in the study by Wright and Sladden (2003) and for a partial own-gender bias in favor of women (Lewin & Herlitz, 2002).

McKelvie (1981) reviewed results of face identification and recognition studies and conducted a number of studies on his own. He concluded that women tended to perform on a higher level in both kinds of face studies, with a same-gender effect for both men and women, but women still tended to identify or recognize more female faces. However, McKelvie (1981) highlighted that his conclusions were not at all fully established, at that time. McKelvie replicated the findings of a female superiority in recognition of other women’s faces. He speculated that a same-gender advantage for women could be grounded in differential interest among women and men to faces so that women would have better memory for other women’s faces, due to differences in socialization. In line with this, another explanation to the female advantage in recognition of same-gender faces are differential interest and that women may be more familiar with other women’s faces (Lewin & Herlitz, 2002). If women have more former knowledge of female faces, then this may facilitate recognition of new women’s faces. Based on the findings and explanations of an own-race bias, Lewin and Herlitz (2002) proposed that familiarity may also be the case for women’s higher memory ability for other women’s faces.

Own-bias effects of both race and gender were found in a study by Rehnman and Herlitz (2006). The participants, Swedish children in ages between 8-10 years, were shown faces of Swedish and Bangladeshi children aged 7-10 years and adults in the ages 20-40 years. They found that it was easier to recognize Swedish faces than Bangladeshi faces. Furthermore, it was found that girls recognized more Swedish and Bangladeshi female
faces than boys did. No gender difference was found for recognition of Swedish male faces, but girls recognized more male Bangladeshi faces than boys did. Rehnman and Herlitz (2006) concluded from their results that the viewer’s familiarity with the ethnicity of the face seemed to have an impact on the magnitude of a gender difference in recognition.

It could be the case of a female predisposition to be more interested in faces as objects than men are in general, in combination with the own-bias, resulting in a very strong own-bias for women. If the proposed reason for the gender difference in face recognition performance is differential interest in faces, with women being more interested in other women’s faces, then one could expect women to remember more female faces if a female face and a male face are shown together at the same time. The reason why this could be expected is that women are thought to be more interested in other women’s faces, and thus it is assumed that women would allocate their attentional resources in first hand on the female faces, resulting in a deeper encoding of these, in comparison to the male faces. So, if women and men differ on their allocation of interest in men and women, then the altering of presentation mode may influence the number of recognized female and male faces, in comparison to the results of the standard presentation mode.

One way of testing this hypothesis would be to manipulate the presentation mode, and present two (a female and a male) faces at the same time in the presentation phase. Performance on the altered face recognition task can be compared to that of the standard, one face at a time, face recognition task. Further, the role of this altered presentation mode in face recognition of same-and other-sex people has not been assessed before. Previous studies have indicated the role of attention and interest in face recognition, therefore this procedure might be a way of study the proposed larger female interest in other women’s faces.

There could be other factors than the presentation mode and interest hypothesis proposed above, which may explain the general female advantage in face recognition. It might be the case for own-gender bias as for own-race bias with the mediating role of contact with the target group, in this case, same- and other-sex persons. Therefore, another proposed explanation for the strong own-gender bias seen in women is that women, compared to men, develop more intimate relationships with their female friends. The same effect might account for higher face recognition for male faces in men with closer same-sex friendships.

*Gender Differences and Friendship*
Adolescents develop friendships in order to fulfill the needs of intimacy, that is, feelings of for example love and security (Sullivan, 1953, in Johnson, 2004). Johnson (2004) reviewed findings of gender differences in friendships: firstly, young women grow more intimate friendships; secondly, they report that it is of importance to keep a relationship intimate; and thirdly, they expect to have more intimate friendships than young men do. There are also tendencies for other-sex friendships to grow, on average, more intimate the older adolescents get (Buhrmester & Furman, 1987).
Friendship has been measured in a variety of ways with many different self-reporting scales. Johnson (2004) conducted a study on gender differences in adolescents' self-reported friendship intimacy. The results of this study showed that girls or young women perceived their friendships to be more intimate (emotionally close) than did boys or young men. The female adolescents reported that they spent more time with their friends and in Grade 8-10 mostly with same-sex friends. According to Johnson (2004), girls are more probable to develop relationships higher in intimacy and emotional closeness levels, in relation to the ones boys’ develop.

Another measure of friendship has been created by Furman and Buhrmester (1985), the Network of Relationships Inventory (NRI). The NRI consists of ten scales, each measuring different aspects of relationships, for example Intimacy, Companionship, Affection, and Conflict. The NRI can be used to assess relations to family members, friends of same- and other-sex, and significant others. From a compilation of scales, a social support measure can be derived. This was done in a study by Furman and Buhrmester (1985). They found that intimacy, affection, and enhancement of worth scales, which are regarded as important in more mature relationships according to the authors, were seen to a larger extent in girls’ self reported relationships with friends than in boys’.

Intimacy in same-sex friendships has been studied in relation to the personality variables (Zarbatany, Conley, & Pepper, 2004). Positive correlations have been found with femininity (Narus & Fischer, 1982, in Zarbatany et al., 2004) and negative have been found with masculinity in preadolescence children (see Jones & Dembo, 1989, in Zarbatany et al., 2004). Zarbatany and colleagues (2004) were interested in individual differences in same-sex friendship needs, from the view of differences in personality traits. They isolated two needs, or traits: communion (need for interpersonal closeness, intimacy, and connection) and agency (individuality, self-expansion, and self-determination). The participants were preadolescents (11-12 year olds) and young adults (19 year olds). It was found that girls and women reported that they wanted mostly communal support in friendships, and they received more of this than did the boys and young men. The male participants reported that they needed both communal and agentic provisions from close same-sex friends. Girls have also been found to search for more intimate friendships earlier than boys do, and that girls same-sex friends’ importance as intimacy providers increase in preadolescence (Buhrmester & Furman, 1987).

In short, gender differences in close relationships have been found. Girls and young women have reported that they spend more time with their same-sex friends (Johnson, 2004), and they report more emotional closeness or intimacy in these relationships (Furman & Buhrmester, 1985; Johnson, 2004) than boys or young men have reported. Grade of qualitative and quantitative contact have been found to mediate the own-race bias (Meissner & Brigham, 2001). Could degree of contact with friends also mediate overall face recognition performance for both men and women? Contact with other people may perhaps be defined as closeness in friendships. One way of measuring degree of depth of friendship could be through using the Network of Relationships Inventory. As women tend to report closer same-sex friendships and also tend to be better at
recognizing other women’s faces, it might be fruitful to assess the possible relation between same-gender face recognition and same-sex friendships. It could be speculated that people who have generally deeper social contact with others could be better in face recognition. In addition to this assumption, more close friendships with same-sex friends may be related to better recognition of same-gender faces. Deeper other-sex friendships could be related to other-gender face recognition.

The first aim of this study was to test the alleged proposition that women are generally more interested in females than males, resulting in a strong own-gender bias. This was assessed through altering the standard presentation mode. It was also of interest to study the men’s performance pattern, if they also would show own-gender bias. In order to investigate this, 41 young women and 16 young men completed two face recognition tasks in a within-subjects design, to compare face recognition memory performance between the two presentation modes. Based on earlier research it was hypothesized that women would outperform men in the recognition of faces in general. It was also hypothesized that women would recognize more female faces if they had to divide their attention between a female and a male face, compared to if female and male faces were presented one and one randomly intermixed.

A second aim of the study was to assess the relation between face recognition and depth of friendships. The reported depth of friendships with same- and other-sex persons may be related to face recognition memory for both men and women. Therefore, it was expected that people who generally had more deep social relations would be better than people with less deep relations on face recognition in general, and that women and men who had more close relationships with same-sex friends would recognize more same-sex persons than those who had less close same-sex relationships. It was also of interest to see if depth of other-sex friendships was related to higher face recognition of other-sex faces, for both men and women.

To test this, the participants in addition to the face recognition tasks completed the Network of Relationships Inventory for same- and other-sex friendships. It was hypothesized that women would report deeper same-sex friendships. In addition, it was hypothesized that deeper friendships overall would be related to better recognition of faces.

**Method**

**Participants**

A number of 57 young adults participated in the study. Their mean age was 18.25 years ($SD = .74$). Of these participants were 41 women, their mean age was 18.17 years ($SD = .64$), and 16 men, whose mean age was 18.34 ($SD = .99$). The participants gave before the test session their informed consent, and the test sessions were during class time. Most participants had finished one or two years in Swedish high-school or the International Baccalaureate program ($M = 1.39$ years, $SD = .62$). The women’s mean value in finished school years was 1.34 ($SD = .43$), and for the men 1.50 years ($SD = .82$).
Material

Face recognition. Two sets of color photographs of faces were included in the face recognition task. The persons’ faces were free from facial hair, jewelry, and glasses, with neutral face expressions, dressed in same-colored t-shirts and standing in front of the same-colored neutral background. Each set (A and B) included 16 female and 16 male faces, a total of 32 faces. Set A and set B were both constructed in two additional versions in order to manipulate the presentation mode. In mode one, faces were presented one at a time for two seconds (1A and 1B), and in mode two one female face and one male face were presented on the screen at the same time during four seconds (2A and 2B). All faces were shown with a PC projector. After an interval of approximately 12 minutes the participants were asked to do a yes-no forced recognition task. This task included 32 faces in each version, of which 16 were targets, that previously had been shown, and 16 were distractors, which were randomly intermixed. The faces were in view, one at a time, for five seconds.

Synonyms. A 25-item multiple-choice test (Nilsson, Bäckman, Erngrund, Nyberg, et al., 1997) was used. The synonyms test contained two practice items and each item consisted of the word to find the synonym for, and five words to choose the target synonym from. The total time for the synonyms test was five minutes. Typically, no gender difference is found on this kind of semantic memory task (Maitland et al., 2004).

Fluency. The participants were instructed to write down as many Swedish words as possible beginning with the letter A and S, respectively. The time given was one minute for each letter. One point was awarded for each correct word. Women tend to perform on a higher level than men on verbal fluency tasks (see for example Maitland et al., 2004).

Mental rotation. The modified version of Shepard-Metzler’s Mental Rotation Test by Vandenberg (1971) for group administration was used to assess mental rotation ability. The test was a paper-and-pencil test and contained two practice items and 10 test items. Each item consisted of one target figure on the left side and four figures on the right side. Two of the four rotated figures were the same as the target figure and the task was to mark these. Two correct answers on each item scored two points; one correct answer scored one point. One or two wrong answers gave zero points for the whole item. The instructions clarified that it would not be of any advantage to guess if the participant had no idea which answer was correct. The participants were given a time limit of six minutes. Men usually outperform women on this task (Voyer et al., 1995).

Friendship. The Network of Relationships Inventory, NRI (Furman & Buhrmester, 1985), was administered to assess grade of close friendship in the participants’ same-sex and other-sex friendships. The NRI consists of 30 questions, contributing to 10 relationship qualities. It was translated to Swedish and shortened to include questions on two relationships; with best same-sex friend and best other-sex friend. The questions concerned how much a certain relationship quality occurred in each of the relationships. The participants were asked questions such as, “How much do you talk about everything with this person?” The rating was done on five-point Likert scales (1- Little or None; 2- Somewhat; 3- Very Much; 4- Extremely Much; 5- The Most). The translation did not
fully work for all anchor points to be exactly the same on all scales as in the original NRI (with an exception for Relative Power which in the original language did differ from the other scales. Relative Power: 1- S/he always does; 2- S/he often does; 3- About the same; 4- I often do; 5- I always do), but the questionnaire was as comparable to the original version as possible. Three items for each friendship quality were averaged to obtain the scale score of it. The qualities included were companionship, conflict, instrumental aid, antagonism, intimacy, nurturance, affection, admiration, relative power, and reliable alliance. To derive a total friendship score for mean depth in social relations, the means from all scales but relative power, which did not fit in the total measure as the anchor points on this item did not give any qualitative information of friendship depth, were averaged. This friendship mean average was computed firstly for friendships overall, and then for same-sex and other-sex friendships respectively.

Procedure
Two groups of participants were tested at two different occasions in classrooms at the participants’ school during normal class time. The participants were not assigned to the two testing sessions in any controlled way. They belonged to two different classes of high-school students and were tested when each class could find the time for it. The time taken to complete all tasks was approximately one hour. The order of tests were as follows: (a) the first set of faces; (b) word comprehension test; (c) mental rotation test; (d) first recognition task; (e) the second presentation of faces; (f) verbal fluency test; (g) NRI; and (h) second recognition task. One group (n = 39; 28 women and 11 men) first viewed the presentation mode one-face first, and the other group (n = 18; 13 women and 5 men) viewed the presentation mode two-faces first. In the second presentation the group which had started to view mode one-face viewed mode two, and vice versa for the other group. Both groups viewed the same set of faces in the same order, but the mode the faces were presented in differed between the groups.

Results
To begin with, independent samples t tests revealed no gender differences in age (t(55) = .829, p = .411) or years of education (t(55) = .866, p = .390) between men and women. Secondly, possible gender differences in cognitive abilities performance were assessed to see if the participants performed as expected, and that reliable further comparisons between the groups of men and women could be done. To assess this, means were computed and tested with independent samples t tests. Significant expected gender differences were found (see Table 1), favoring women on the fluency test (t(55) = 2.181, p = .033), and favoring men on the mental rotation test (t(55) = -3.321, p = .002). No significant gender difference was found on the synonym test (t(55) = 1.917, p = .060).
Table 1. Means for men and women on the fluency, mental rotation, and synonym tasks.

<table>
<thead>
<tr>
<th></th>
<th>Women M (SD)</th>
<th>Men M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>15.35 (2.79)*</td>
<td>13.53 (2.95)</td>
</tr>
<tr>
<td>Mental Rotation</td>
<td>9.73 (5.41)*</td>
<td>14.88 (4.82)</td>
</tr>
<tr>
<td>Synonyms</td>
<td>19.93 (2.97)</td>
<td>17.69 (3.11)</td>
</tr>
</tbody>
</table>

(*p < .05)

Results of Face Presentation Mode

In order to assess the effect of presentation mode a 2 (sex of viewer: female, male) x 2 (gender of face: female, male) x 2 (presentation mode: one-face, two-faces) ANOVA was computed. Sex of viewer was the between-subject variable. Gender of faces (female and male) and the presentation modes (one-face and two-faces) were the within-subject variables. The dependent variable was the number of correctly identified faces (e.g. hits minus false alarms). The result showed no main effect of sex of participant ($F(1, 55) = 2.680, p = .107$) meaning that women and men did not differ significantly in their general face recognition performance. No main effect of presentation mode was found ($F(1, 55) = .183, p = .670$), indicating that there was no difference in performance between mode one or mode two, and no main effect of the gender of the seen face ($F(1, 55) = .057, p = .812$), showing overall no effect on participants’ performance of the gender of the presented faces. Further, the expected interaction effect between gender of face and sex of participants was also non-significant ($F(1, 55) = .970, p = .329$) meaning that women were overall not better on female faces. Men and women were equally good at recognizing faces in different presentation modes, as indicated by the non-significant interaction effect between presentation mode and sex of participant ($F(1, 55) = 1.761, p = .190$). A significant effect of gender of face and of presentation mode ($F(1, 55) = 4.208, p = .045$) indicated that men were easier to recognize in the one-face mode, and women in the two-faces mode. The tree-way interaction was not significant ($F(1, 55) = .946, p = .335$), showing that there was no effect of an interaction between gender of face, presentation mode, and sex of participant.

Follow-up t tests for independent samples showed that women did not significantly outperform men on total face score ($t(55) = 1.637, p = .107, d = .47$), or on male faces ($t(55) = .858, p = .395, d = .27$), but outperformed men in recognition of female faces ($t(55) = 1.972, p = .054, d = .55$).

To further investigate the effect of mode and gender of face, separate ANOVAS for men and women were computed. The dependent variable was number of faces recognized (e.g. hits minus false alarms), and the within-subject factors were gender of face (female and male) and mode (one-face and two-faces), see Table 2.
Table 2. Mean (SD) number of recognized faces (Hits - False alarms) for men and women on face gender and presentation mode. The maximum score in each condition was 8.

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mode 1</td>
<td>Mode 2</td>
</tr>
<tr>
<td>Female faces</td>
<td>4.00(1.90)</td>
<td>4.56(2.19)</td>
</tr>
<tr>
<td>Male faces</td>
<td>4.50(2.07)</td>
<td>4.38(2.00)</td>
</tr>
<tr>
<td>Total</td>
<td>8.50(2.92)</td>
<td>8.94(3.28)</td>
</tr>
</tbody>
</table>

For women no significant main effect of gender of face was found ($F(1, 55) = 1.303, p = .260$), meaning that women were equally good in recognition of female and male faces overall. No significant main effect of presentation mode was found ($F(1, 55) = 3.186, p = .082$). Importantly, and as hypothesized, an interaction effect was found between presentation mode and gender of face ($F(1, 55) = 8.477, p = .006$) showing that women performed on a higher level on female faces when a male and a female face was presented together and on male faces when they were presented one at a time (see Figure 1).

![Figure 1](#)

Figure 1. Mean number of recognized faces (Hits - False Alarms) for women.

For men, no main effects of gender of face ($F(1, 15) = .207, p = .656$) was found, meaning that men performed equally well no matter face gender. Further, no main effect of presentation mode ($F(1, 15) = .205, p = .657$) was found. Men tended to perform equally well if the faces were presented one and one as when faces were presented together. No interaction effect of gender of face and presentation mode was found ($F(1, 15) = .369, p = .553$), showing that men performed on the same level across presentation modes and gender of faces (see Figure 2).
Figure 2. Mean number of recognized faces (Hits - False Alarms) for men.

Thereafter, as there were specific hypothesis, simple effects were evaluated with $t$ tests for paired samples for men and women separately. Women did not significantly recognize more female faces when seen in mode two than in mode one ($t(40) = 1.289, p = .205, d = -.30$). Further, women recognized significantly more male faces in the mode where faces were shown one at a time than in mode two ($t(40) = 3.467, p = .001, d = .71$).

The results showed a similar pattern for the men as the one seen in women, but the tendencies were non-significant. No difference for the men was found between recognition of women in mode one or mode two ($t(15) = .872, p = .397, d = -.27$), and for between male faces between the different modes ($t(15) = .150, p = .882, d = .06$).

**Results of Face Recognition and Depth of Friendships**

In order to investigate hypothesized relationships between depth of friendships and performance on the face recognition tasks, firstly gender differences generally in friendship qualities were assessed with independent samples $t$ tests. Significant gender differences, with women reporting higher grades of these qualities, were found in the same-sex friendship overall score ($t(54) = 2.337, p = .023$) and in the same-sex friendship qualities affection ($t(55) = 2.132, p = .037$), companionship ($t(55) = 2.239, p = .024$), intimacy ($t(55) = 2.851, p = .006$), and nurturance ($t(55) = 2.958, p = .005$).

To investigate if there were any differences between same- and other-sex friendships qualities for men and women separately, $t$ tests for paired samples were computed. Women consistently reported higher level of affection ($t(39) = 3.415, p = .002$), companionship ($t(39) = 3.783, p < .001$), instrumental aid ($t(38) = 2.884, p = .006$), intimacy ($t(39) = 5.330, p < .001$), nurturance ($t(39) = 4.350, p < .001$), and reliable alliance ($t(39) = 5.143, p < .001$) in their same-sex friendships than in other-sex friendships (see Table 3). In addition to this, the women’s composite same-sex friendship overall score was significantly higher than the other-sex friendship score ($t(37) = 4.589, p$
As for men, no significant differences were found between reported grades of friendship qualities in same- and other-sex friendship scales.

Table 3. The mean values and standard deviations for same- and other-sex mean friendship quality score. Maximum score on each scale was 5.

<table>
<thead>
<tr>
<th></th>
<th>Men Same-sex</th>
<th>Men Other-sex</th>
<th>Women Same-sex</th>
<th>Women Other-sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affection</td>
<td>3.21(.93)</td>
<td>3.06(.98)</td>
<td>3.77(.88)*</td>
<td>3.23(.80)</td>
</tr>
<tr>
<td>Companionship</td>
<td>2.79(.91)</td>
<td>2.50(.67)</td>
<td>3.41(.80)**</td>
<td>2.74(.83)</td>
</tr>
<tr>
<td>Instrumental Aid</td>
<td>2.52(1.05)</td>
<td>2.25(.76)</td>
<td>2.82(.84)*</td>
<td>2.42(.81)</td>
</tr>
<tr>
<td>Intimacy</td>
<td>2.71(1.13)</td>
<td>2.67(1.26)</td>
<td>3.61(1.05)**</td>
<td>2.42(1.15)</td>
</tr>
<tr>
<td>Nurturance</td>
<td>2.60(.90)</td>
<td>2.92(1.09)</td>
<td>3.43(.97)**</td>
<td>2.74(1.00)</td>
</tr>
<tr>
<td>Reliable Alliance</td>
<td>3.27(1.20)</td>
<td>3.17(1.29)</td>
<td>3.93(.1.12)**</td>
<td>2.79(1.16)</td>
</tr>
<tr>
<td>Friendship Overall</td>
<td>3.27(.56)</td>
<td>3.18(.61)</td>
<td>3.67(.60)**</td>
<td>3.16(.61)</td>
</tr>
</tbody>
</table>

(* p < .05, ** p < .001)

To study the hypothesized relationships between memory for faces and depth of friendships, correlations were computed with Pearson’s r for mean total friendship score and total face recognition score. No significant correlation was found between these variables \(r(52) = .178, p = .197\). To further explore the possibility of any relationship, the reported grade of friendship for men and women, respectively, were correlated to total face recognition score. No significant relationship was found for women \(r(36) = .094, p = .575\) or men \(r(14) = .257, p = .338\).

Firstly, correlations between overall same-sex friendship score and women’s recognition of female faces, and men’s overall same-sex friendship score in relation to their recognition of male faces, were computed to see if overall same-sex friendship score was related to same-sex recognition score. Recognition for women of female faces and same-sex friendships overall score were not related \(r(38) = -.008, p = .963\). The men’s recognition of male faces were not either related to same-sex friendships \(r(14) = -.188, p = .485\).

Secondly, other-sex friendship scores were correlated to women and men’s recognition of other-sex faces. No relationships were found between women’s recognition of male faces and other-sex friendships \(r(36) = .069, p = .683\) or men’s recognition of female faces and other-sex friendships \(r(14) = .318, p = .230\).

**Discussion**

The first aim of the present study was to investigate the possibility of women being more interested in females than in males, through comparing performance on the standard presentation mode, one face at a time, with recognition of the altered presentation in
which faces were presented two at a time. As hypothesized, women tended to recognize more female faces when the faces were presented two and two. However, the result was only significant when both female and male faces were considered. It was non-significant when simple effects were computed and the female faces were compared across modes. This might be due to the small sample of participants, as the effect may have been too small for becoming significant with only 41 female participants. The effect size of more male faces being recognized when presented one at a time than two at a time was significant, and larger, with a $d$-value of .71. The $d$-value computed for more female faces recognized in mode two than in mode one was -.30. This result is, in spite of its statistical non-significance and a small $d$-value, of major importance. It is in line with the proposed, but previously unstudied, possible explanation of the strong female own-gender bias to be caused by a larger female interest in other women (Lewin & Herlitz, 2002; Rehnman & Herlitz, 2006). This finding may indicate that women, when a female and a male face are presented at a same time, allocate their attention rather to the female than the male face, as the result was better memory in women for the female faces. This finding could benefit from the explanation of superior encoding as a reason for the own-race bias (Meissner et al., 2005), as it seems that women directed their attention to the female faces in first hand, thus the female faces were better encoded than the male faces, resulting in a better memory for the female faces.

The ecological validity is of value to mention in this context, as the standard face recognition procedure does not seem to have high ecological validity (Wright & Sladden, 2003). It can be argued that we do not watch people one at a time in everyday life and therefore alternative procedures ought to be used to complement results obtained from the standard face recognition procedure. Based on this, it ought to be of value to present more than one face at a time to try to study the face recognition process under more realistic conditions. Therefore the altering of the presentation mode seen in this study could be an additional way of studying human face recognition as new results might emerge.

Men’s performance pattern was also of interest. No own-gender bias for men was found, which is coherent with most previous results where own-biases have been considered (Lewin & Herlitz, 2002; Rehnman & Herlitz, 2006). Men’s performance pattern on the female and male faces across modes was in the same directions as the women’s performance. Both men and women were better at remembering male faces when they were presented one at a time, and female faces when presented two at a time, but the results for the men was consequently non-significant. It seems as if both men and women directed attention toward a female face when two faces are presented together, but women’s performance seem to have been influenced by the altered presentation mode to some higher extent. An unexpected result was that the effect size for men of more female faces being recognized when presented two and two ($d = -.27$). As very few males participated in this study, and no similar hypothesis about a larger male interest in females have been proposed, it remains to see in future studies with similar design if men will show a similar pattern as women are thought to show. If men will recognize more female faces when faces are presented two and two, this could further support an interest
hypothesis, but the reason why males in first hand would direct their attention to a female face must be addressed.

As hypothesized, a general female advantage in face recognition was found in this study across different modes and gender of faces, which is in line with results from previous face recognition studies (Herlitz et al., 1997; Lewin et al., 2001; Rehnman & Herlitz, 2006). This general female advantage in face recognition may have a biological origin, supported by results from studies of infants. Girls have been found to look for a longer time than boys on a face (Connellan et al., 2000).

The expected own-gender effect for women was not significant in this study, which is at odds with previous findings (Lewin & Herlitz, 2002; McKelvie, 1981; Rehnman & Herlitz, 2006). This interaction effect has been found to be large (Lewin & Herlitz, 2002; Rehnman & Herlitz, 2006). The reasons for this could be that the female faces may have been more homogeneous looking than the male faces, or that the male faces overall were more heterogeneous looking, which might have resulted in that the male faces were more easily recognized and thus the interaction effect became non-significant for female faces overall. The effect of perceived distinctiveness is according to Meissner et al. (2005) well documented in face recognition. Faces rated as distinctive or atypical were more easily recognized than faces rated as usual (e.g. Brigham, 1990, in Meissner et al., 2005). The distinctiveness effect may have influenced the recognition of the male faces in this study. The faces which were showed in the presentation modes included 16 faces in each, and each recognition task contained a total of 32 faces. The number of faces included in the presentations and in the recognition tasks in this study was comparably small to that of other studies where at least forty faces, or sixty faces, are used in the recognition tasks (Lewin et al., 2001; Lewin & Herlitz, 2002; Rehnman & Herlitz, 2006). If it was the effect of distinctiveness that influenced the recognition of male faces, this could possibly have been reduced if a larger number of target and distractor faces were used.

The second aim of the study was to assess the hypothesized role of deep friendships in face recognition. The expected gender differences were found in same-sex friendship qualities, for instance in intimacy where young women have been found to report higher levels than young men (Furman & Buhrmester, 1985; Johnson, 2004), but no significant relationships were found between friendship overall score and face recognition in general. Further analysis did not reveal any relationships between same-sex depth of friendship and same-gender face recognition, or between other-sex friendships and recognition of other-gender faces. However, the correlations between men’s recognition of faces in general and depth of friendships and between men’s recognition of female faces and depth of other-sex friendships indicated a relation in the proposed direction. As the results were non-significant, no conclusions can be drawn on the role of friendship in face recognition. It could be speculated that the Network of Relationships Inventory was not an appropriate measure of the construct friendship as intended, and another operationalization of friendship ought to be used to measure grade of contact and/or interest in others. This could be assessed in future studies, with more effort laid on grade of contact with same- and other-sex friends.
This study had some limitations. To begin with, the number of participants was small, which partially might have reduced the possibility of generalizing the result. Secondly, the sample might not have been representative of the population of young adults, as it could have differed in some respects. Thirdly, two classes of students were tested and they were not comparable in group size. This meant that one larger group was tested with one procedure and fewer were tested with the second procedure. Ideally the two groups had been equally sized. In addition, the number of women who participated was two thirds of the total number of participants, meaning that the number of men was comparably smaller. This could also be argued to influence the possibility to generalize, as it is possible that the men’s face recognition performance differed from the average young male performance. On the other hand, as the expected gender differences in mental rotation and verbal fluency were found, the sample was small, but did not seem to deviate from the normal population in other aspects. However, the difference in performance on the synonyms test was close to significance, and women tended to perform on a higher level than men, which is contrary to the expected lack of a difference in performance. A significant difference in self-reported interest before the tests were taken was found, and women reported a significantly higher interest to do a word-comprehension test than men did, which might to a small extent have accounted for the observed difference in performance.

Finally, to mention the study’s strengths, the performance of women and men were mainly studied within respective group. Therefore the different sized groups of men and women are not of great importance and ought not to be considered as a limitation. Further, in this study a within-group design was used. Both groups viewed the same sets of faces, and the presentation mode was the only condition that was altered. In this way, the participants became their own controls, thus improving the reliability in the present study.

In conclusion, no support was found for any relationships between depth of friendship and face recognition performance, not for friendship depth in general and face recognition overall, and not for same-sex depth of friendships and men and women’s separate recognition of same-gender faces. The results indicated a general female advantage in face recognition performance, and some support was found for a greater interest in women for females than males. This finding is of importance, as the sets of faces seemed to have included distinct looking male faces which could have been easier to recognize, and in spite of this, women tended to recognize more female faces when the faces were presented two and two. A performance pattern that was noticeable in the men’s performance as well. Taken together, the results from this study supported previous findings showing a general female advantage in face recognition, and that this in combination with women’s larger interest in female faces than male faces may result in the female own-gender bias. The present study can be seen as a pilot for assessment of the role of interest in face recognition. It has implications for future face recognition research, as interest and attention in earlier research has been proposed to contribute to the own-race bias (Meissner et al., 2005), and also to account for the strong own-gender effect seen in women, in combination with a general female in face recognition (Lewin & Herlitz, 2002; Rehnman & Herlitz, 2006). The results ought to be regarded with some
caution due to the small sample of participants, which to a degree may limit the possibility to generalize the results. It is therefore highly important to conduct a larger study in the future of the interest hypothesis, as it may partly explain the strong own-gender bias found in women.

References


