WHERE ARE KIN RECOGNITION SIGNALS IN THE HUMAN FACE?

MARIA F. DAL MARTELLO         LAURENCE T. MALONEY

Dipartimento di Psicologia Generale
Università di Padova

Department of Psychology
Center for Neural Science
New York University

M. F. Dal Martello: Maria.DalMartello@unipd.it
http://dpg.psy.unipd.it/sch_docenti.php?id=42

L. T. Maloney: Laurence.Maloney@nyu.edu
http://www.psych.nyu.edu/maloney/index.html


Corresponding author:

Maria F. Dal Martello
Department of General Psychology
University of Padova
via Venezia 8
Padova
Italy

Tel: +39 049 658112
Email: maria.dalmartello@unipd.it
ABSTRACT

We report two experiments intended to determine where in the face the cues signalling kinship fall. In both experiments, participants were shown thirty pairs of photographs of children’s faces. Half of the pairs portrayed siblings and half did not. The 220 participants were asked to judge whether each pair of photos portrayed siblings. We measured the effect on kin recognition performance of masks that covered the upper half of the face or the lower half (Experiment 1) and the eye region or the mouth region (Experiment 2). In Experiment 1 we found that the signal detection estimate of performance $d'$ decreased only 5.3% (n.s.) when the lower face was masked but by more than 65% when the upper face was masked. We tested whether combination of kinship information from the two halves of the face can be treated as optimal combination of independent cues and found that it could be. In Experiment 2, we found that masking the eye region led to only a 20% reduction (n.s.) in performance while masking the mouth region led to a non-significant increase in performance. We also found that the eye region contains only slightly more information about kinship than the upper half of the face outside of the eye region.
INTRODUCTION

Kin recognition is thought to play an important role in the social organization of species (De Bruine, 2002, 2004, 2005; Platek, Raines, Gallup et al, 2004; Platek, Keenan & Mohamed, 2005). The “inclusive fitness” theory of Hamilton (1964a, 1964b) presupposes that organisms have some ability to recognize their own kin and to discriminate different degrees of relatedness in their relatives (Mateo, 2002). Recognition of kin to self through self referential phenotype matching has been documented in species from bacteria to primates (e.g. Fletcher & Michener, 1987; Hepper, 1991; Chapais & Berman, 2004).

Additionally, some primates are able to recognize closely related individuals of their own species that are not closely related to themselves (Parr & De Waal, 1999; Parr, Winslow, Hopkins & de Waal, 2000). Cheney & Seyfarth (1989, 1999), for example, found that Vervet monkeys, after witnessing a fight between a relative and an individual of another family, will become aggressive towards other members of the second family.

Human observers, also, can detect relatedness between individuals unrelated to themselves: in parents and children pairs (Nesse, Silverman, & Bortz, 1990; Christenfeld & Hill, 1995; Brédart & French, 1999; Bressan & Dal Martello, 2002; Bressan & Grassi, 2004) and in sibling pairs (Maloney & Dal Martello, 2006). Participants, in all of these studies, were well above chance in classifying individuals as near relatives or not, given only photographs of their faces. Humans can even detect relatedness in pairs of individuals of another primate species (Vokey, Rendall, Tangen, Parr & de Waal, 2004).

The primary aim of the study is to determine where the kin signals used by human observers are in the human face. We concentrate on collateral kin recognition, specifically classification of pairs of children as siblings or non-siblings. There are no
other studies (to our knowledge) in the kin recognition literature about the spatial location of kin signals in the face. There are studies on face recognition which look for the relative importance of parts of the face in recognition tasks. Goldstein & Mackenberg (1966) and Fisher & Cox (1975), for example, found that the upper features were more important than the bottom ones for face recognition. It is interesting to consider the hypothesis that the same regions of the face that signal identity would also signal kinship.

Studies on face recognition also indicate that the eyes are a very salient feature for identity recognition. Fisher & Cox (1975) and McKelvie (1976) found that omission of the eyes in a test picture of the face was more disruptive to recognition than omission of the mouth. Roberts & Bruce (1988) found that masking the eyes led to a sharper reduction in performance in a sex recognition task than masking the mouth. Davies, Ellis & Shepherd (1977) investigated the cue salience of internal facial features changing single features of faces constructed with a Photophit Kit. They found that the eyes were more informative for recognition than the mouth. Chimpanzees' performance in a recognition task of conspecifics' faces deteriorated when only the eyes where masked, but did not significantly decrease when only the mouth was masked. On each trial, the primate observers were shown three faces: a sample picture of a face and two comparison pictures of faces one of which portrayed the same individual as portrayed in the sample picture. They were rewarded if they matched the sample picture to the comparison picture that portrayed the same individual (Parr, Winslow, Hopkins & de Waal, 2000).

Keating and Keating (1993) found that rhesus monkeys' recognition of identi-kit built human faces deteriorated only when the eyes or brows were altered. Recognition was not affected by alteration or removal of the other internal features of the face. Measures of fixation patterns showed that the monkeys concentrated attention mostly on
the eyes. Kyes and Candland (1987) found a similar visual preference for the eyes when baboons viewed the features of a conspecific’s face.

Our stimuli include pairs of children that are close in age and pairs that are widely separated in age. Given normal facial development in children, these age differences complicate the observer’s task. Even if all the traits in the human face are equally and highly hereditable as Kohn (1991) concludes, the time course of development is radically different in different parts of the human face. From the literature on face morphology (Enlow & Hans, 1996) we know that the face below the eyes changes dramatically in size and shape from an age of one year to adulthood (Figure 1).

Figure 1. Development of the Human Skull. The neonate skull is enlarged to emphasize the relative growth of the cranial vault (A) and the lower face (B) (from The John Hopkins Hospital Center for Craniofacial Development & Disorders Web Site).

The lower face becomes longer and larger to accommodate the expansion of the nasal region and the eruption of deciduous and permanent teeth. The cranium, instead, grows much less than the lower part of the face. A neonate’s basicranium is 60 to 65 per cent of its final size, it reaches 90 per cent of its full development when the child is 5 to 7 years old. Most of the growth of the cranial vault, like the brain’s growth, takes place during the first year of life. The size of the face (below the eyes) at birth is around one eight of the total size of the head, while in a young adult it is half of it. We
emphasize that both neurocranial and lower facial dimensions are highly heritable (Kohn, 1991). They differ in the extent to which they change over the course of childhood. It is plausible to expect that facial cues signaling kinship will be found in parts of the face that vary least across the lifespan.

Our first hypothesis, then, was that the cues present in the upper, developmentally stable, part of the face will carry most weight in a kin recognition task consistent with previous results in the face recognition literature. In the first experiment, we tested this hypothesis and related hypotheses by asking adults to perform signal detection tasks where the signal is ‘close genetic relatedness’ between two children. Our stimuli are color pictures of children faces seen from the front. The children’s ages varied from 17 months to 15 years, covering a large part of the developmental period. We measure participants ability to detect ‘close genetic relatedness’ using standard signal detection methods described below. We expect that kin recognition performance will be disrupted more when the upper half of the face is not visible and it is of interest to quantify how much each region of the face contributes to kin recognition.

Our second hypothesis was that the eye region carries many of the cues signaling kinship, consistent with results in the face recognition literature just described, and consistent with the consideration that eye size changes little over the course of post-natal development, reaching 70% of their adult size by birth and then growing rapidly in the first 12 months of life (Gossmann, Mohay & Roberts, 1999). In a second experiment, we will contrast ability to recognize kinship when either the mouth region or eye regions of faces are masked, and compare performance in both of these conditions to a control, ‘Full Face’ condition.

---

1 The term ‘face’ is used in the face morphology and dental literature to refer to the lower part of the face in ordinary usage, from the eyes to the chin. We will use the term ‘lower face’ instead.
As a working definition, we define a signal\(^2\) as any measurement derived from a face (Tversky, 1977) that carries useful information relevant to a visual task. We recognize that there may be signals that fall entirely under one of our masks but that there may be signals that signal valid information about kinship that are not available when, for example, either the upper or the lower half-face is masked. An example of such a signal would be eye separation divided by the distance from the center of the line joining the pupils of the eyes to the mouth (a ‘configurational’ feature of the sort proposed in Rhodes, 1988). Masking either the upper or lower half-face would make it impossible to estimate this signal accurately.

Such a ‘full face signal’ would be important to our analyses here only if it were, first of all, a good cue to kinship (that is, it really is a kin recognition signal) and, second, the observer gave considerable weight to it in assessing kinship.

From this viewpoint, in Experiment 1, there are potentially three classes of kin recognition signals that can be differentiated by how masking affects kin recognition. Those that are disrupted by masking the upper half face but not the lower (‘upper signals’), those that are disrupted by masking the lower half face but not the upper (‘lower signals’) and those that are disrupted by masking either the upper or the lower half faces (‘full face signals’). If there were no full face signals in this sense, we would expect that we could predict performance in the unmasked condition from measured performances in the two masked conditions.

In the first experiment, we will also test whether detection performance in the two masked conditions can be used to predict performance in the unmasked condition using standard methods of signal detection theory. In effect we are testing whether this

\(^2\) The term ‘signal’ here is a synonym for the term ‘feature’ as used by, e.g. Tversky (1977). Some authors drop the requirement that a ‘feature’ must carry objectively useful information relevant to a task. That is, the visual system may mistakenly make use of a ‘feature’ that unfortunately is not a ‘signal’. 
third class of full face signals carries substantial additional information about kinship that is used by observers.
EXPERIMENT 1

Methods

Participants. 109 people, recruited in public places at the University of Padova, were alternately assigned to one of three conditions: Full Face (FF), Lower Half Masked (LHM) or Upper Half Masked (UHM) as described below. There were 18 males and 16 females in Condition FF, 16 males and 19 females in Condition LHM, and 19 males and 20 females in Condition UHM. Their ages ranged from 19 to 36 years (median age 22 years).

Photographic Material. 72 color photographs, each depicting a child from the neck to the top of the head, were used. The pictures had been taken by the experimenters or their assistants under controlled lighting conditions. Of the 72 children depicted in the pictures, half were girls and half were boys. We used Adobe Photoshop® to obliterate all background detail, replacing it by a uniform dark grey field (33% of maximum intensity in each of R, G and B channels). The ages of the children ranged from 17 months to 15 years. The distributions of age differences for related and unrelated pairs were matched. The distribution of gender was also counterbalanced across related and unrelated pairs. The facial expressions were neutral or close to neutral. All came from three adjacent provinces of Northern Italy, Padova, Mantova and Vicenza. All were Caucasian in appearance. The parents of each child gave appropriate permission for their child’s photograph to be used in scientific experiments. We asked and received separate parental permission to use the photographs in Fig. 2 and 3 as illustrations here.

There were three conditions in the experiment. In the Full Face (FF) condition the pictures were presented without any alteration so that all the face was visible. In the other two conditions we altered the pictures covering part of them. In the Lower Half
Masked (LHM) condition, we added a trapezoid blue figure to each photograph that covered the lower part of the picture and that left visible the upper half of the face. In the Upper Half Masked (UHM) face condition we covered the upper part of the picture leaving visible the lower half of the face. The line dividing the face in upper and lower part was parallel to the line where the eyes lie and passed through the tip of the nose (Figure 2).

**Figure 2. Experiment 1: Experimental Conditions.** On each trial observers saw a pair of photographs of children’s faces and were asked to judge whether the children were sibling’s or not. There were three conditions, illustrated here: A. The Full Face (FF) Condition (no mask); B. The Lower-Half Masked (LHM) Condition; and C. The Upper-Half Masked (UHM) Condition.

For privacy reasons, we did not verify by DNA fingerprinting that sibling pairs shared two parents. Recent research using DNA fingerprinting shows that the median rate of “extrapair paternity” is much lower than previously thought, under 2% (See Simmons, Firman, Rhodes & Peters, 2004, for review). Consequently, it is unlikely that any large proportion of our siblings share only one parent (mother but not a father). In any case, the presence of half-siblings would have little effect on the outcome of our experiment. Such half-sibling would have 25% of their DNA in common, rather than 50%, but would still be more closely related than non-sibling pairs and their presence in the sample across all conditions should not affect the comparisons across conditions that are central to our analyses.

**Picture Pairs.** Of the 72 photographs, sixty were used in the main experiment. The remaining 12 were used just in the Familiarization and Training phases of the experiment. These phases will be described below. The sixty photographs used in the
main phase of the experiment included 15 pairs of biological siblings and 15 pairs of children who were genetically unrelated. We refer to the pairs in the first group as related and in the second as unrelated. Within each group of 15, five pairs depicted two boys, five pairs depicted two girls and five pairs depicted a boy and a girl. The twelve photographs used in the Familiarization and Training stages included three pairs of biological siblings and three pairs of unrelated children. In the Full Face condition the photographs showed the face fully visible, in the other two conditions the face was partially masked as described above.

Procedure. The experiment was conducted in a computer classroom at the University of Padova. Observers viewed all stimuli on a computer monitor and responded by marking forms provided. The experiment was self-paced and consisted of three phases. 1. Familiarization. The observer was first asked to perform a simple recognition memory task that involved all of the experimental stimuli for their condition. All 72 photographs of faces were shown in groups of six per display in random order. The purpose of this part of the experiment was to familiarize the observer with the range of faces he would see in the main experiment. The observer was asked to study the display and told that, immediately after studying each group, he would be shown a probe photograph and asked to report whether this photograph had been among the group of six just studied. The probe photographs were the non-experimental photographs described above which were not used in the main part of the experiment. 2. Training. The observer practiced the response for his condition (either Full Face or Upper Half Masked or Lower Half Masked) on six pairs of photographs that did not overlap with the photographs used in the main part of the experiment. These pairs were drawn from the non-experimental photographs organized so that there were three pairs that were biological siblings and three that were not. The purpose of this part of the experiment was simply to let the observer become comfortable with the procedure and response required. In this phase the observers were
told that half the pairs portrayed genetic siblings and were asked to classify each pair as related or not related. 3. Main. The observers then were told that half pairs of the set of photographs in the main phase were of genetic siblings and half were of unrelated children. Their task, as in the training phase, was to classify each pair as related or not related. The 30 pairs of photographs were presented in random order. We used two separate randomizations and observers were assigned one of them at random.

**Results and Discussion**

Signal detection estimates of sensitivity $d'$ and likelihood criterion $\beta$ were used to measure performance in the different conditions (Green & Swets, 1966/1974). These values are reported in Table 1. We first discuss the $d'$ values which are also summarized in Figure 3.

<table>
<thead>
<tr>
<th>Condition</th>
<th>$d'$</th>
<th>SD($d'$)</th>
<th>$\beta$</th>
<th>SD($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF</td>
<td>1.187</td>
<td>0.085</td>
<td>0.885</td>
<td>0.046</td>
</tr>
<tr>
<td>UHM</td>
<td>0.408</td>
<td>0.075</td>
<td>0.927</td>
<td>0.020</td>
</tr>
<tr>
<td>LHM</td>
<td>1.124</td>
<td>0.081</td>
<td>0.893</td>
<td>0.042</td>
</tr>
</tbody>
</table>

**Table 1. Results for Experiment 1 by Condition.** The $d'$ estimate and the likelihood criterion $\beta$ for the signal detection analysis. Standard deviations were estimated by a bootstrap procedure (Efron & Tibshirani, 1993) based on 10,000 replications.
Figure 3. Experiment 1: Results. The $d'$ measures for the three experimental conditions are plotted as a bar graph. The confidence intervals shown to 95% confidence intervals were estimated using a bootstrap procedure (Efron & Tibshirani, 1993) with 10,000 replications. The $d'$ values for the FF (Full Face) condition and the LHM (Lower-Half Masked) condition are significantly greater ($p < 0.001$) than that for the UHM (Upper-Half Masked) condition.

In carrying out the hypothesis tests below, it would be appropriate to correct for multiple tests by a Bonferroni correction. However, for all of the tests in this experiment and the following, a Bonferroni correction would not change any conclusions as the reader can verify. Consequently, we simply present the p-values (or bounds on the p-values) for each test.

A $d'$ value of 0 corresponds to chance performance and a $d'$ value of 3.5 a practically perfect performance. The $d'$ value in the FF condition was significantly different from 0 ($z = 14.838; p < .0001$, one-tailed). The one-tailed test is justified as it is plausible to assume $d' \geq 0$. The participants, when able to observe all the face, could classify the pairs as siblings or not siblings markedly above chance level. The $d'$ values are similar to those we have found in earlier, related work (Maloney & Dal Martello, 2006). The $d'$ value in the LHM condition was also significantly different from 0 ($z = 6.486; p < .0001$, one-tailed). When the lower half of the face was hidden, the observers could still discriminate well between siblings and non-siblings pairs. Performance in FF and in LHM condition did not differ significantly ($z = 0.553; p = .580$, two-tailed). The absence of information from the lower half face resulted in a small decrease in sensitivity that was not statistically significant.

The $d'$ value in the UHM condition was significantly different from 0 ($z = 5.440; p < .0001$; one-tailed): participants can classify the pairs at a level above chance even if all the upper half of the face is hidden. It appears that there are useful signals for kin recognition in the lower half face as well. Performance in the UHM condition ($d' = 0.41$)
was significantly worse than either the FF condition (z = -7.104; \( p < .0001 \)) or the UHM condition (z = -6.486 \( p < .0001 \)). Even if it was still possible for the observers to detect relatedness at a level above chance when just the bottom half of the face was visible, the performance in this condition deteriorated markedly when compared to the other two conditions.

We emphasize that, when we fail to reject a null hypothesis, it would not be correct to conclude that the null hypothesis is true or even that we have evidence in favor of the null hypothesis. This is the common fallacy of “accepting the null hypothesis” (See Loftus, 1996). When, for example, we do not reject the null hypothesis that the \( d' \) for LHM was equal to that for the FF, we have not shown that the \( d' \)’s are equal or that the lower-half face adds nothing to performance in judging relatedness. We have simply shown that we cannot reliably measure the difference with our experimental design: the difference is correspondingly small and the best information available for the magnitude of the difference is the estimates of \( d' \) in Table 1 and the accompanying estimates of their standard deviations.

If the information available in the upper and lower half face were statistically independent and if there were no additional sources of information available only in the full face condition (the ‘full face signals’ discussed in the introduction), then we would be able to predict the \( d' \) for the full face from the \( d' \) values in the other two conditions (Green & Swets, 1966/1974):

\[
d_{FF}' = \sqrt{(d_{UHM}')^2 + (d_{LHM}')^2}
\]

(1)

where \( d_{FF}', d_{UHM}', d_{LHM}' \) are the \( d' \) values in the three corresponding conditions. If we form a prediction of \( d_{FF}' \) based on the values for the other two conditions in Table 1, we find that this prediction \( \hat{d}_{FF}' = 1.196 \), is remarkably close to the value observed,
$d'_{FF} = 1.187$, and the two values are not significantly different ($p > 0.05$). Thus, although the $d'$ for the FF condition was not significantly different from that for the LHM condition, the small difference observed is consistent with combination of statistically independent information from the two half faces.

In conclusion, there are useful cues to kinship available when either the upper or lower half of the face is visible, but the upper half of the face contains almost as much useful information as the full face. The results are also consistent with the claim that statistically independent information from the two half faces is combined in the FF condition and that the contribution of full face signals in collateral kin recognition is negligible. We return to this point in the general discussion.

The $\beta$ values reported in Table 1 measure the bias of the response towards classifying as kin or not kin. The $\beta$ values for each condition are not significantly different from one another ($p > 0.05$). That is, observers did not show any change in bias because of the presence or absence of the masks. Most importantly, they do not become more cautious (adopting a stricter criterion) when just half of a face is visible.

Note that the three conditions showed a common bias. We tested the $\beta$ values tested against 1 and in all conditions, the difference from 1 was significant ($p < 0.01$) in all cases. The actual prior odds are 1:1 that the pairs are related (half the pairs portray siblings) and the observers were given this information. Still the observers were slightly biased in favor of classifying the pairs as related. That is, they err in the direction of misclassifying unrelated pairs as related (Type 1 Error).
EXPERIMENT 2

Introduction

In this experiment we tried to determine the weight of relatedness cues present in the eyes and in the mouth. As we described above, studies on face recognition indicate that the eyes are more important than the mouth for identity recognition in human and other primate species.

The literature on face morphology and development describes the eyes area as relatively invariant during development (Enlow & Hans, 1996). We hypothesize that by analogy with the identity recognition studies results the eyes should contain more useful kin recognition cues than the mouth and consequently that masking the eye region should lead to a marked drop in performance in kin recognition. We test our hypothesis measuring kin recognition performance in an experiment with the same design, stimuli and task used in the first experiment but using smaller masks covering smaller regions. The masks were sufficient to cover the eyes in one experimental condition and the mouth in the other. The face was fully visible in the control condition. The three conditions are illustrated in Figure 4.

Figure 4. Experiment 2: Experimental Conditions. On each trial observers saw a pair of photographs of children’s faces and were asked to judge whether the children were sibling’s or not. There were three conditions, illustrated here: A. The Full Face (FF) presentation (no mask); B. The Eyes-Masked (EM) Condition; C. The Mouth-Masked (MM) Condition.
Methods

Participants. One hundred and eleven people were recruited in public places or personally contacted and randomly assigned to three conditions: Full Face (FF), Eyes Masked (EM), and Mouth Masked (MM). Each containing 37 observers (13 males and 24 females). Their ages ranged from 20 to 31 years (median age 23).

Material. The stimuli were the same 30 pairs of photographs as those used in the main phase of Experiment 1. In the FF Condition the photographs were presented without alterations, in the other two conditions we altered the pictures by adding masks. In the Mouth Masked condition (MM), we added a blue rectangle to each photograph that just covered the lips of the child. In the Eyes Masked condition (EM) we added the smallest blue rectangle that just covered the eyes and eyebrows of the child. We asked and received parental permission to use the photographs in Figs. 4 and 5 as illustrations here.

Procedure. The experiment, like the first experiment, was conducted in a computer classroom at the University of Padova. Observers viewed all stimuli on a computer monitor and responded by marking forms provided. The experiment was self-paced. The observers were told that half pairs of the set of photographs in the main phase were of genetic siblings and half were of unrelated children. Their task was to classify each pair as related or not related. The 30 pairs of photographs were presented in random order.

Results and discussion

Signal detection analysis was used. We report the results for each condition as $d'$ and $\beta$ values (Table 2). We first discuss the $d'$ values which are also summarized in Figure 5.
Table 2. Results for Experiment 2 by Condition. The $d'$ estimate and the likelihood criterion $\beta$ for the signal detection analysis. Standard deviations were estimated by a bootstrap procedure (Efron & Tibshirani, 1993) based on 10,000 replications.

<table>
<thead>
<tr>
<th></th>
<th>$d'$</th>
<th>SD($d'$)</th>
<th>$\beta$</th>
<th>SD($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF</td>
<td>1.023</td>
<td>0.080</td>
<td>0.971</td>
<td>0.041</td>
</tr>
<tr>
<td>EM</td>
<td>0.818</td>
<td>0.078</td>
<td>1.043</td>
<td>0.033</td>
</tr>
<tr>
<td>MM</td>
<td>1.107</td>
<td>0.080</td>
<td>0.960</td>
<td>0.043</td>
</tr>
</tbody>
</table>

The estimated $d'$ values in all three conditions were significantly greater than 0 (FF: $z = 12.787$, $p < 0.0001$; EM: $z = 1.487$, $p < 0.0001$; MM: $z = 13.838$: $p < 0.0001$, all one-tailed). Performance in the masked eyes condition was significantly worse ($z = 2.587$; $p < 0.01$, 2-tailed) than performance in the mouth masked condition. This outcome parallels results in the recognition literature (Goldstein & Mackenberg, 1966; Fisher & Cox, 1975). The group of observers who could see all the face except for the eyes did worst in the kin recognition task than observers who could see all the face but the mouth. The differences between the full face conditions and the two masked conditions were not significantly different from 0 (FF vs MM: $z=1.835$; $p = .458$, two-tailed; FF vs EM: $z=1.835$; $p = .067$).
Figure 5. Experiment 2: Results. The $d'$ values for the three experimental conditions are plotted as a bar graph. The confidence intervals shown to 95% confidence intervals were estimated using a bootstrap procedure (Efron & Tibshirani, 1993) with 10,000 replications. The $d'$ for the FF (Full Face) condition is not significantly different ($p > 0.05$) from the $d'$ values for the other two conditions.

Consider next the pattern of the outcomes of the hypothesis tests in Experiment 2. For convenience we will use the designator of each Experimental condition to denote the $d'$ estimate of that condition. We failed to reject the hypotheses $d'_{FF} = d'_{EM}$ and $d'_{FF} = d'_{MM}$ while we did reject $d'_{EM} = d'_{MM}$. Yet, there is no contradiction here as there would be if we thought we had shown $d'_{FF} = d'_{EM}$ and $d'_{FF} = d'_{MM}$ but $d'_{EM} < d'_{MM}$. As we noted above, failing to reject a point null hypothesis $AB = 0$ does not imply that $A$ really is equal to $B$ (Cohen, 1994; Loftus, 1996).

But we can say more. One reviewer advanced the following argument: if we accept that it is plausible that $d'_{MM} < d'_{FF}$ (masking the mouth can only reduce performance) then the pattern of results $d'_{EM} < d'_{MM}$ (the outcome of an hypothesis test) and $d'_{MM} < d'_{FF}$ (our assumption) would entail the conclusion that $d'_{EM} < d'_{MM}$. We note that this conclusion is based on the assumption $d'_{MM} < d'_{FF}$ as well as our statistical results and, while plausible, goes beyond the normal range of statistical argument.

The estimates of effect sizes in Table 2 together with the estimated standard deviations of these effects are the best estimates we have concerning the quantity of kin recognition information available in different parts of the face. We will analyze them further together with the results summarized in Table 1 in the General Discussion.

The $\beta$ values in the three conditions are not significantly different from each other. They are also not significantly different from 1 (We omit detailed reports of the z-values).
Since the prior odds that the children are related are 1:1, the participants are correctly using them, they do not show any bias in favor of classifying them as kin rather than not. We note that participants did not have the *familiarization* and *training* phase in Experiment 2 and perhaps this difference in procedure accounts for the bias found in Experiment 1 but not in Experiment 2.
GENERAL DISCUSSION

We reported two experiments intended to determine what regions of the face contain the cues signalling kinship. In both experiments, participants were shown thirty pairs of photographs of children’s faces in random order. Half of the pairs portrayed siblings and half did not. A total of 220 participants were asked to judge whether each pair of photos portrayed siblings (yes-no task).

Pairs of photographs were either presented with both faces fully visible (the unmasked condition) or with opaque masks covering corresponding parts of both faces. In Experiment 1, there were three conditions, one unmasked and two masked. In the first masked condition, masks covered the lower half of the faces of both children in each pair. In the second, masks covered the upper half of the faces. Observers participated in only one of the conditions. In Experiment 2 there were also three conditions, one unmasked and two masked. In the first masked condition, the masks covered a small region containing the eyes. In the second, the masks covered a small region containing the mouth. We analyzed the data using standard signal detection methods (Green & Swets, 1966/1974).

In Experiment 1 we found that performance did not deteriorate significantly when the bottom half of the face was masked. The estimated $d'$ with the bottom half of the face masked was only 5.3% less than that for the full face. Performance, instead, deteriorated significantly (about 20%) when the upper half face was masked. This result is consistent with previous results in the face recognition literature (Goldstein & Mackenberg, 1966; Fisher & Cox, 1975) and consistent with what we know about the development of the face (Kohn, 1991; Enlow & Hans, 1996). We conclude that the observer extracts considerably more information about kinship from the upper half of the child’s face than from the lower.
We also found that observers are well above chance in judging kinship when the upper half of the face is masked, although \( d' \) in this case was significantly lower than when the lower half of the face was masked (a decrease in percentage terms of more than 65% in \( d' \) value).

We also examined whether performance in the full face condition was appreciably better than would be predicted by combining performance in the two masked conditions (Green & Swets, 1966/1974) assuming optimal cue combination (Landy, Maloney, Johnston & Young, 1995). This outcome would be expected if, for example, ratios of separations of signals that fell in both regions (e.g. eye separation divided by eye to mouth separation) were good cues to kinship and the observer gave considerable weight to them in assessing kinship. We found that, in this case, the ‘whole’ was the ‘sum of the parts’. There was no indication that such full face signals were used.

In Experiment 2, we found that masking the eye region led to only a 20% reduction in performance as measured by \( d' \). This reduction was not statistically significant (reported above) and we would certainly not conclude that the eye region dominates kin recognition performance given that masking of the eye region leads to such a small (and non-significant) reduction. But we approach this issue (the relative importance of the eye region to kin recognition) in a different way.

In Experiment 1, we examined the additivity of information across the lower and upper halves of the faces. We next consider a second additive decomposition where we partition the face in the lower half (LH), the eye region (E) and the upper half excluding the eye regions (UH\(\backslash\)E). The results of this decomposition will give us a better idea of the information about kinship contained in the eye region and in the upper part of the face outside of the eye region. We use the identity

\[
(d_{FF}')^2 = (d_E')^2 + (d_{EM}')^2
\]  

(2)
based on partitioning the full face into the eye region and its complement. We use a second identity

\[(d'_{FF})^2 = (d'_{LH})^2 + (d'_{E})^2 + (d'_{UH,E})^2\]  \hspace{1cm} (3)

based on partitioning the full face into lower half, eye region, and the part of the upper face not in the eye region. We have an estimate of \(d'_{EM}\) from Experiment 2, and an estimate of \(d'_{LH}\) from Experiment 1 (it is just the estimate of \(d'_{UHM}\)). We have estimates of \(d'_{FF}\) from both experiments and we use the average of these estimates. We can then solve the equations above for estimates of \(d'_E = 0.743 \pm 0.165\) and \(d'_{UH,E} = 0.709 \pm 0.102\). The values following the ± are bootstrap estimates of the standard deviation (Efron & Tibshirani, 1993). The estimate \(d'_{LH} = 0.408 \pm 0.075\) for the lower half of the face (Table 1) completes the decomposition of the face into three regions. We emphasize that this computation depends on the assumption that information about kinship in non-overlapping regions of the face is independent and correctly described by Eqs. 2, 3 above.

If we accept these estimates, it is evident that the eye region contains only slightly more information about kinship (as measured by \(d' = 0.743\)) than the upper half of the face (\(d' = 0.709\)) outside of the eye region and the difference is not significant (\(z = 0.1753; p = 0.861\), two-tailed). Our results contrast with typical conclusions in the face recognition literature (Fisher & Cox, 1975; McKelvie, 1976; Roberts & Bruce, 1988; Davies, Ellis & Shepherd, 1977). Yet, given differences in experimental design, stimuli, and methods between the experiments reported here and experiments in the face recognition literature just cited, we can draw no firm conclusions. We do, however, advance the conjecture that the eye region contains primarily information about the specific identity of the individual, not their kin group. This conjecture is intriguing (and unexpected) and certainly deserves further study. Experiments are needed to determine
the relative importance of the eye region in kin recognition tasks and in identity recognition tasks using comparable stimuli and methods.

The picture that emerges from this study is intriguing. We can summarize the information in a face as a list of signals (measurements made on the face). Some signals are of use in assessing age and gender and these signals are correspondingly of little use in assessing kinship. Signals of recognition can include those that signal age, gender and kinship, but except in the case of identical twins, we expect that there are also signals specific to the individual, the basis for recognition. Our results combined with previous results in the literature, indicate that the visual system selects and combines whatever signals are available in un-occluded parts of the face according to task. Visual performance is robust in that large parts of a face can be masked and still the visual system can extract task-relevant information as in the UHM condition of Experiment 1.

By comparing performance across task with selective masking of regions of the face, we can develop a better understanding of what these signals are and how they enter into different visual tasks involving faces.
ACKNOWLEDGMENTS

LTM was supported by NIH grant EY08266, MFDM was supported by funds from the Italian Ministero dell'Università e della Ricerca Scientifica e Tecnologica. We acknowledge Davide Bianconi, Francesca Constantini and Carla Scaglierini for assistance in collection of data and taking of photographs and an anonymous reviewer for suggestions.
REFERENCES


Loftus, G. R. (1996), Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5*, 161-171.


