*Note: The following manuscript has been published in Cognition, which can be openly accessed at:* <https://www.sciencedirect.com/science/article/pii/S0010027725001155>

How Much Face Identity Information Is Required for Face Recognition?

Mintao Zhao1 and Isabelle Bülthoff2

1 School of Psychology, University of East Anglia, United Kingdom

2 Department of Human Perception, Cognition, and Action,

Max Planck Institute for Biological Cybernetics, Germany

Declarations of interest: None

Correspondence concerning this article should be addressed to Mintao Zhao (email: mintao.zhao@uea.ac.uk), School of Psychology, University of East Anglia, Norwich, NR4 7TJ United Kingdom, or to Isabelle Bülthoff (email: isabelle.buelthoff @tuebingen.mpg.de), Department of Human Perception, Cognition, and Action, Max Planck Institute for Biological Cybernetics, D-72076 Tübingen, Germany.

Abstract

Many studies have shown that degradation of face identity information impairs face recognition, however, when such information degradation reaches the limit of our face recognition ability remains unclear. Here we systematically decreased face identity information by morphing an increasing number of faces together and investigated how much identity information is required for recognizing a face in a morph. Our results show that participants could identify half of faces mixed in 3-identity morphs using only their memory of these faces (Experiment 1) and, when perceptual information is available, they could recognize two of three faces mixed in a morph (Experiment 2). When we systematically reduced the contribution of each identity to a face morph from 50% to 6.25% (i.e., morphing 2 to 16 faces together; Experiments 3 and 4), participants could still consistently recognize faces in a morph containing as little as 12.5% of their identity information. Moreover, familiarity with faces enhanced participants’ performance, whether they were asked to recognize all faces mixed in a morph in one go (Experiments 1 and 2) or to recognize them individually (Experiments 3 and 4). Finally, image-based similarity between the faces and morphs could predict how decreasing identity information impairs face recognition performance. Together, these results not only help quantify the minimum information required for face recognition but also offer new insights into the representational differences between familiar and unfamiliar faces.

*Keywords*: face recognition, face perception, identity information, face familiarity, morphing

**Highlights**

• People can recognize multiple face identities mixed equally in a face morph

• 12.5% identity information in a morph is sufficient for above-chance recognition

• Image similarity predicts correct recognition of multiple faces in a morph

• Familiarity enhances face recognition with limited identity information

How Much Face Identity Information Is Required for Face Recognition?

**1 Introduction**

An image of a single face can carry information about multiple identities (e.g., images obtained by morphing two faces together; Jacques & Rossion, 2006; Rotshtein et al., 2005). When the faces of two people are morphed into one, the resulting morph can be considered as a weakened representation of those two face identities, as only half of their diagnostic features are preserved. How well people can recognize faces depicting weakened identity information has been extensively investigated to understand the mechanisms of face recognition and to improve the detection of fraudulent use of identity documents (e.g., Robertson, Kramer, & Burton, 2017). For instance, identity information in face images can be degraded by blurring or pixelating images (Bachmann, 1991; Costen, Parker, & Craw, 1996; Gilad-Gutnick, Yovel, & Sinha, 2012), by reducing idiosyncratic information through face morphing (Jiang, Blanz, & O’Toole, 2009; Leopold et al., 2001), and by disrupting pigmentation, reflectance, and texture of facial images (Bruce & Langton, 1994; Gilad, Meng, & Sinha, 2009; Russell et al., 2007). While these studies elegantly demonstrate that degrading face identity information impairs our ability to recognize faces, they have rarely investigated at which level of degradation face identification fails. To address this question, in the present study, we systematically mixed an increasing number of faces into one face morph and tested how our ability to recognize a face is affected by gradually decreased identity information available in a morph. Our results may shed light on one fundamental question of face recognition research: what is the minimum identity information required to give a face its identity (e.g., Bachmann, 1991)?

Previous studies have shown that people can recognize a face or differentiate between two faces based on a fraction of identity information. For example, Leopold et al. (2001) have shown that, for newly learned faces, people need about 11% of identity information (i.e., at 11% of the distance from the average face to the original face identity in a hypothetical multi-dimensional face space, Valentine, 1991) to reach a 50% identification accuracy, and people showed near perfect recognition performance with only one-third of identity information. Wilson and colleagues (Gao & Wilson, 2013; Wilson, Loffler, & Wilkinson, 2002) tested how well people can discriminate between two unfamiliar synthetic faces by varying the geometric information between them. They found that participants could reach a 75% accuracy when the two faces differed in terms of 6–8% of their geometric information (e.g., distance between the original faces and the average face). Recently, we tested when people lost their ability to differentiate between two morphed faces created with an equal number of faces. When the image-based difference between both morphs was at least 25% of the average differences between any two individual faces, discrimination performance was above chance level; when the ratio dropped to or below 16%, performance was at chance level (Bülthoff & Zhao, 2021). Therefore, a small proportion (e.g., 6-16% as shown above) of identity information appears to be sufficient for face recognition and face discrimination.

When two faces are morphed into one image, people seem to need a high proportion of identity information to recognize them. For instance, using a two-alternative-forced-choice (2AFC) face classification task (i.e., judging if a morph displays person A or B), Beale and colleagues (Beale & Keil, 1995; Levin & Beale, 2000) showed that a 40% or higher proportion of face information is needed for its recognition. People are generally unable to identify a face in a morph when the face contributes 20% or less information to the morph. Similar results have been frequently observed in other studies using 2AFC tasks (Campanella et al., 2003; Chauhan et al., 2020; McKone, Martini, & Nakayama, 2001; Kikutani, Roberson, & Hanley, 2008, 2010; Jacques & Rossion, 2006). These findings support the categorical perception of face identity. That is, a linear increase of identity information in a morph leads to a nonlinear steplike increase in its identification, in the shape of a sigmoid function. For a two-identity morph, 50% or higher information of a target identity is usually required to go across the categorical boundaries, leading to its recognition.

The difference in sensitivity between face identification (i.e., classifying a face as belonging to a learned identity) and face discrimination (i.e., judging if two faces are the same or different) was also observed in previous studies (Beale & Keil, 1995; Levin & Beale, 2000; Kikutani et al., 2008, 2010; McKone et al., 2001). For instance, McKone et al (2001) found that, for a 2AFC face classification task, people exhibit categorical perception of face identity. Again, 40% or more of face identity information in a morph was required to identify that face. In contrast, for a face discrimination task (e.g., determining which of two faces is more like person A), people needed only a small difference in identity information between two faces (e.g., 6-14%; see also Wilson et al., 2002). These results suggest that when two faces are morphed together, a higher proportion of identity information is required to identify a face in morph than to differentiate between two morphs.

Why do people need a high proportion (e.g., ≥40%) of identity information to recognize a face in a two-identity-morph? One possibility is that, for a familiar face (e.g., famous or newly learned faces), its representation in face space (Valentine, 1991) may have a small “attractor field”, where only faces depicting a sufficiently high amount of its identity information can fall within its “attractor field”, which could function as a gating system that determines whether a stimuli can activate memory about the face and trigger its recognition (Tanaka et al., 1998; see also Loffler et al., 2005). When two face are morphed together, although people can differentiate one morph (e.g., 10% of face A with 90% of face B) from another (e.g., 30% of face A with 70% of face B), these two morphs may still not be recognized as face A because both are outside of its “attractor field” in face space (e.g., Beale & Keil, 1995; McKone et al., 2001).

Alternatively, the amount of identity information needed for recognition may be overestimated with the 2AFC face classification task. Following the above example, the correct choice of identity A might be hindered when the competing identity B is perceptually more dominant. That is, people may perceive both face identities in a morph but are forced to report the dominant one given the task demand. Hence, the observed higher proportion of identity information may reflect the minimum information required for winning the identity competition rather than for above-chance recognition. Consistent with this hypothesis, when tested using a task without such identity competition, face identification can be achieved with much less identity information (e.g., 11%, Leopold et al., 2001; see also Robertson et al., 2017). Similarly, people can still recognize faces when identity information is dramatically degraded by blurring or pixelating of facial images (e.g., 16-20 pixels per face horizontally; Bachmann, 1991; Bindemann et al., 2013; Costen et al., 1996; Lander, Bruce, & Hill, 2001; Ramon et al., 2015). These results suggest that the “attractor field” of familiar faces in face space may be larger than that estimated using 2AFC tasks, and images that depict only a small proportion of familiar identity information may suffice for activating face memory and triggering recognition.

In the present study, we investigated how much identity information is required for above-chance face identification by systematically varying the amount of identity information displayed in a face morph. In Experiments 1 and 2, we created morphs with three faces (i.e., each contributing 33.3% of identity information) and asked participants to recognize them based on their memory of these faces (Experiment 1) or on both memory and perceptual information (Experiment 2). In contrast to the 2AFC face classification task, where participants must pinpoint a unique identity in a face morph, our task allowed participants to report multiple identities without the needs to suppress non-dominant competing faces.

In Experiments 3 and 4, participants saw a face morph next to a test face, which could be a face used to create the face morph or not. Participants were asked to judge whether the test face was used to create the face morph. If the tasks used in Experiments 1 and 2 can be understood as analogous to the whole-report task in Sperling (1960), where participants were asked to recognize all faces mixed in a morph, the task used in Experiments 3 and 4 is analogous to his part-report task, which examined participants’ ability to recognize individual faces in a morph. By systematically reducing the amount of information contributed by individual faces (e.g., by morphing an increasing number of faces together), this task provided a sensitive measure of minimal idiosyncratic facial information required for above-chance identification.

Our second goal was to examine whether familiarity with faces modulates the minimum identity information required for their identification. Differences between recognition of familiar and unfamiliar faces have been well documented in the literature (Hancock, Bruce, & Burton, 2000; Johnston & Edmonds, 2009; Megreya & Burton, 2006; Young & Burton, 2018). In comparison to unfamiliar faces, familiar faces are argued to be more robustly represented and recognized (particularly under non-ideal conditions), with a culmination for personally familiar faces like family members, friends, colleagues, and oneself (Natu & O'Toole, 2011; Ramon & Gobbini, 2018; Tong & Nakayama, 1999). While it is intuitive to assume that people would need less information for identifying familiar than unfamiliar faces, such difference has not been clearly established. Some studies suggest that this is the case (e.g., Ramon et al., 2015), whereas other studies suggest that the reverse may be true (e.g., Chauhan et al., 2020). For a face that contributes 40% of information to a face morph, people rarely recognize the face in more than 20% of the time in a 2AFC face classification task (cf. face discrimination task in Beale & Keil, 1995), regardless of whether the face morphs are created with famous faces (Beale & Keil, 1995; Rotshtein et al., 2005), with newly learned faces (Levin & Beale, 2000; McKone et al, 2001), or with novel faces without any prior learning (Jaques & Rossion, 2006).

To examine if familiarity enhances face identification when only a fraction of identity information is available (i.e., in a face morph), we tested two groups of participants. One group had no visual experience with the faces used to create the morphs (i.e., unfamiliar group), whereas the other group was personally familiar with those faces (i.e., familiar group). This manipulation allowed us to test whether familiarity with faces helps people recognize more faces in a morph (Experiments 1 and 2) and require less identity information for face recognition (Experiments 3 and 4). The results may inform us whether familiar and unfamiliar faces differ in terms of the minimum identity information required to carry their identities.

**2 Experiment 1**

Experiment 1 investigated whether people can recognize multiple face identities from a single facial image. Face morphs mixing two faces have been used to test how familiarity, identity, sex, or race of faces is perceived (Busey, 1998; Bülthoff & Newell, 2004; Bülthoff, Manno, & Zhao, 2023; Levin, 1996; Levin & Beale, 2000; Rossion, 2002; Rotshtein et al., 2015). However, whether, and if so, how well people can recognize multiple face identities mixed in a morph remains largely unexplored. To address this question, we created morphs by equally mixing three faces and asked participants to identify these faces in a morph. This task allowed participants to report all recognized identities without having to suppress non-dominant competing ones. Importantly, the amount of information contributed by each face (33.3%) is below what is required for their recognition in a 2AFC face classification task (i.e., 40% or more). If people need 40% or more identity information to recognize a face, as shown in previous studies (e.g., Beale & Keil, 1995; McKone et al., 2001), participants should have difficulty recognizing any of the three faces. In contrast, if previous 2AFC classification tasks overestimated the minimum information required for face identification, one-third of identity information may be sufficient for recognizing all three faces (e.g., Leopold et al., 2001; Wilson et al., 2002). Following the convention of previous research (e.g., Busey, 1998, Tanaka et al., 1998), hereafter, we refer to the faces/identities used to create face morphs as “parent face” or “parent identity”. This helps our description of procedures and results and does not imply that recognition of faces in a morph share the same processes as kin recognition.

## 2.1 Method

### 2.1.1 Participants

Thirty-one colleagues (12 females, age was not asked individually but ranged roughly between 25 and 60 years) from our institution participated in the study. Power analysis using G\*Power 3.1 indicated that for a one-way repeated ANOVA (3 measurements), to achieve a statistical power (1-*β*) of 0.80 with a medium effect size (*f* =0.25) at *α* =.05, 28 participants are required.

Our participants were familiar with the colleagues whose faces were used to create the stimuli (i.e., target faces). Participation was voluntary and no reimbursement was offered. Note that “familiar” means that our colleagues knew the colleagues whose faces were used as parent faces, and all had been at our institute for six months or longer. Exactly how long they had been familiar with each other and how frequently they encountered each other were not assessed.

### 2.1.2 Stimuli

Face morphs were created using ten male and ten female faces of our colleagues (ethnicity: White; facial expression: neutral). The methods to create face morphs were similar to those used in previous studies (Bülthoff & Zhao, 2020, 2021; Jiang et al., 2009). We first 3D-scanned these colleagues’ faces and then created face morphs using the 3D Morphable Model (Blanz & Vetter, 1999; Troje & Bülthoff, 1996). All scanned faces were in correspondence to each other, with each face represented by 70,000 points indexing facial shape and texture. To create a 3-identity morph, we averaged the shape and texture values of three parent faces at each point and rendered the resulting average face (morph). Note that this morphing process is carried out on the 3D faces, not 2D images. We created thirty 3-identity morphs (ten female, ten male, and ten mixed-sex morphs; **Figure 1**). For mixed-sex morphs, five were created using two female faces and one male face and another five were created using two male faces and one female face.

A close up of a person's face

Description automatically generated

**Figure 1**. Example of female morphs used in Experiment 1. F01, F04, F08 and so on indicate which parent faces were used to create the morph (e.g., F01 refers to the female parent identity No. 1, this information was not shown to participants). Note that the three morphs shown here are visually different although they share some parent identities (e.g., F04 was used to create all three morphs).

### 2.1.3 Procedure

We first accustomed our participants to the appearance of the 20 parent faces, which were shown individually with their names for 15 s. Participants were informed that all test morphs were created using three of those twenty faces. At test, participants saw the 30 morphs, one at a time, and their task was to identify which colleagues’ faces were used to create the morphs. The trials were grouped into six blocks of five trials. At the beginning of each block, participants were informed whether the following face morphs were male, female, or mixed-sex morphs. Each morph was shown for 15 s, followed by a blank screen, and participants could start to write down their answers upon seeing the morph. The order of blocks (female-male-mixed-female-male mixed) and trials was the same for all participants.

Due to colleagues’ availability, we tested twenty-six of them together in a large room, with face morphs presented on a projection screen (3 to 5 m from participants’ seating position). Another five participants performed the same task individually sitting in front of a computer monitor. On their answer sheet, one page listed the names of the 20 colleagues, another page showed the trial number (i.e., 1-30, this information was shown next to the test morph on the screen) followed by three response boxes to enter the names of the parent faces they identified from the morph. Participants could leave any number of response boxes empty if they could not identify all three parent faces.

## 2.2 Results and Discussion

On average, participants accurately recognized 1.5 parent identities from a morph (**Figure 2**). A repeated measures ANOVA revealed that performance differed across the three types of face morphs, *F*(2,60) = 15.60, *p* < .001, *ηp*2 = .342. Paired *t*-tests showed that recognition of male morphs (1.68 ± 0.08) was better than that for female morphs (1.45 ± 0.09) and mixed morphs (1.35 ± 0.08), both *t*(30)> 3.52, *p* ≤ .001, Cohen’s *d* > 0.63; whereas the latter two showed no difference, *t*(30) = 1.44, *p* = .161, Cohen’s *d* = 0.26 (hereafter, unless noted otherwise, all post hoc comparisons reported in this paper survive the Bonferroni correction). If our participants had randomly chosen three names for each trial, the chance level performance would be 0.45 correctly guessed identities (see **Appendix A** for details). One-sample *t*-test confirmed that participants’ performance was significantly higher than chance level for each morph condition, all *t*(30)> 10.68, *p* < .001, Cohen’s *d* ≥ 1.92. Participants were able to read out more than one identity for each type of face morph, all *t*(30) ≥ 4.51, *p* < .001, Cohen’s *d* ≥ 0.81.

A diagram of different colored shapes

Description automatically generated

**Figure 2**. Recognition performance for male, female, and mixed-sex morphs in Experiment 1. Solid dots, individual performance; open dots, mean performance. Error bars, standard errors of the mean (SEM). Dashed line, chance level at 0.45.

These results demonstrate that the 3-identity morphs depicted more than one recognizable face identity to people who were familiar with them. Participants recognized one-half of all parent identities using their memory about them. In contrast to the 2AFC face classification task (e.g., Beale & Keil, 1995; Jacques & Rossion, 2006; McKone, Martini, & Nakaya, 2001), in which high contribution of identity information (e.g., 40% or more) is required for recognition, our results indicate that people can recognize multiple familiar faces based on a smaller proportion of identity information (e.g., 33.3%).

Nevertheless, our participants rarely recognized all three parent faces in a morph although they were personally familiar with those people. One possibility is that one-third of identity information is not sufficient for signaling familiarity for all parent identities. For instance, Rossion (2002) showed that when a face morph consists of 40% of a newly learned face and 60% of an unfamiliar face, participants only recognized the morph as familiar faces in 26.5% of trials. When information of familiar face identity dropped to 20%, performance further reduced to 12.5% of trials (see also Beale & Keil, 1995; Levin & Beale, 2000). Alternatively, perceptual differences between what is memorized about a familiar face in real life and what is shown in the face morph (i.e., 3D scanned face without hair and makeup) might limit participants’ ability to recognize familiar faces. Although we showed individual scanned parent faces at the beginning of the experiment, participants may make judgments based on their memory of colleagues’ faces rather than based on the memory of those 3D-scanned faces. In Experiment 2, we tested whether providing perceptual cues helps participants recognize multiple faces in a morph.

**3 Experiment 2**

Experiment 2 further investigated how well people can recognize multiple face identities in a face morph. Instead of asking participants to recognize faces based exclusively on memory (Experiment 1), we showed the same 20 parent faces as in Experiment 1 alongside the test morph and asked participants to identify which of them were used to create the morph. This manipulation allowed us to test if perceptual information from the parent faces facilitates their recognition. Again, if 40% or more identity information is required for recognition, presenting parent faces together with the test morph should not enhance participants’ performance compared to Experiment 1, as neither Experiment 1 nor 2 showed this amount of face identity information in a morph.

Importantly, the task used in Experiment 2 does not require participants to be familiar with the parent faces (e.g., people may perform the task using an image-matching strategy), which allowed us to explore whether familiarity enhances recognition of multiple identities from one face image. We tested two groups of participants, one group had no prior visual experience with the parent faces (i.e., unfamiliar group), whereas the other group were our colleagues who were personally familiar with the parent identities (i.e., familiar group). If both the familiar and unfamiliar groups adopted the same strategy to perform the task (e.g., image-matching; Hancock et al., 2000; Young & Burton, 2018), we would expect equivalent performance between both groups. If familiarity enhances face representation, whether qualitatively (e.g., switching from image-based to abstract representation) or quantitatively (e.g., enlarging its “attractor field” in face space), performance should be better for the familiar than the unfamiliar group.

## 3.1 Method

### 3.1.1 Participants

For the unfamiliar group, 26 participants took part in this study (18 females; aged 19 to 55 years, mean age, 27.9). None of them was familiar with the parent identities used in this study. Written informed consent was obtained before the experiment. They received small monetary payment for their participation. For the familiar group, as most colleagues had participated in Experiment 1, we only managed to test seven colleagues (4 females, mean age 37.0, range: 28-59 years) who were familiar with the parent identities but had not participated in Experiment 1. Their participation was voluntary, and no reimbursement was offered.

As in Experiment 1, power analysis indicated that for a 2 by 3 mixed design (with 1-β = 0.80, *f* = 0.25, α = .05), 28 participants in total are required for testing within-subject factor or within-between interaction, and 86 participants are required for testing between factors. Therefore, our sample size was appropriate to test the effect of face morph condition and its interaction with familiarity group, however, caution should be taken when interpreting the effect of between-group factors (e.g., familiarity) given our small sample size.

### 3.1.2 Stimuli

Stimuli were the same as in Experiment 1. For each trial, a face morph was shown at the center of the screen (about 9°x14°), surrounded by the twenty parent faces (each subtended about 1.3°x1.9°, **Figure 3**). The placement of the parent faces remained identical across all trials for all participants, with female faces on the upper and the right side and male faces on the lower and left side.

A screenshot of a person's face

Description automatically generated

**Figure 3**. Example of a trial in Experiment 2. The central 3-identity morph is surrounded by all 20 potential parent faces. The three parent faces used to create the morph are indicated with a red frame. Other parent faces are replaced with an identity-neutral face for illustration purposes due to permission issues.

### 3.1.3 Procedure

The study was conducted using the E-prime software (Psychological Software Tools, Inc., Pittsburgh, PA, USA). Participants had two practice trials before completing thirty experimental trials. As in Experiment 1, before each block of 5 trials, participants were informed about whether the following face morphs were female, male, or mixed-sex morphs. Participants were asked to identify the three parent faces used to create the morph by clicking on individual target faces. One click would make a face red-framed (indicating that it is selected, **Figure 3**) and clicking on it again would remove the selection (and red frame). An error message would appear if participants attempted to select more than three faces. When participants were certain about their choices, they clicked the “next” button to proceed to the next trial. Each trial was displayed until a response was made. To reduce forced guesses, participants were allowed to proceed to the next trial even if they had chosen less than three faces. With this procedure, the unfamiliar and familiar groups chose three parent faces in 80% and 87% of total trials, respectively, and they chose two or three parent faces in 97% and 100% of total trials, respectively.

## 3.2 Results

Mean numbers of correctly recognized faces are shown in **Figure 4**. A 2 (participants group: familiar vs. unfamiliar) × 3 (morph type: female, male, mixed) mixed ANOVA revealed that the familiar group recognized more identities (2.31 ± 0.10) than the unfamiliar group (1.79 ± 0.05), *F*(1,31) = 20.45, *p* < .001, *ηp*2 = .40. Performance also differed significantly across the three types of morphs, *F*(2,62) = 9.62, *p* < .001, *ηp*2 = .24. The interaction between familiarity and morph type was not significant, *F*(2,62) = 2.16, *p* = .124, *ηp*2 = .07. Paired t-test revealed that recognizing faces from the mixed-sex morphs (1.71 ± 0.08) was more difficult than from male or female morphs (2.06 ± 0.07 and 1.92 ± 0.05 respectively), both *t*(32) > 4.10, *p* < .001, Cohen’s *d* > 0.71. The difference between male and female morph conditions was significant but did not survive the Bonferroni correction, *t*(32)=2.20, *p* = .035, Cohen’s *d* = 0.38.

Both groups were able to recognize multiple identities in one morph. For the familiar group, the number of recognized identities was equal to or higher than two, all *t*(6) >2.36, *p* ≤ .056, Cohen’s *d* > 0.89. For the unfamiliar group, they were able to recognize two identities from male morphs, *t*(25) = 0.61, *p* = .551, Cohen’s *d* = 0.12, but recognized less than two identities for female and mixed morphs, both *t*(25) ≥3.29, *p* ≤ .003, Cohen’s *d* > 0.64. Again, performance was above chance level for all conditions, all *t* ≥14.48, *p* ≤ .001, Cohen’s *d* > 2.75.

A group of different colored shapes

Description automatically generated with medium confidence

**Figure 4**. Recognition performance on male, female, and mixed-sex morphs in Experiment 2. Solid dots, individual performance; open dots, mean performance. Error bars, SEM. Dashed line, chance level at 0.45.

To test whether presenting parent faces together with the morph improves performance, we combined data from both familiar groups in Experiments 1 and 2 and performed a 2 (condition: presence vs absence of parent faces) × 3 (morph type: female, male, mixed) mixed ANOVA. Participants showed better performance with parent faces present (2.31 ± 0.15) than absent (1.49 ± 0.07), showing a significant main effect of condition, *F*(1,36) = 24.22, *p* < .001, *ηp*2 = .402. The main effect of morph type was significant, *F*(2,72) = 8.01, *p* < .001, *ηp*2 = .182. The interaction between condition and morph type was not significant, *F* < 1. Paired t-test showed that the overall performance was better for male morphs than for female or mixed morphs, *t*(37) > 3.95, *p* < .001, Cohen’s *d* > 0.64, and the latter two conditions showed no significant difference, *t*(37) = 1.47, *p* = .149, Cohen’s *d* = 0.24.

Given the small sample size for the familiar group, we further evaluated the effect of between-group factors (e.g., familiarity or experiments) using Bayesian ANOVAs (van den Bergh et al., 2020; see **Appendix B** for full results). Consistent with the above results, the Bayesian analysis showed strong evidence supporting that familiarity had an effect on recognition performance (BFInclusion = 316.895) and that the differences between Experiments 1 and 2 had an effect on recognition of familiar faces (BFInclusion = 525.297).

## 3.3 Discussion

When perceptual information about parent identities is available, familiar and unfamiliar groups can recognize 2.3 and 1.8 identities on average, respectively. This result provides further evidence that one-third of identity information from parent faces can lead to their recognition in a face morph (albeit still imperfectly). Together with Experiment 1, our results suggest that the higher amount of identity information required for face classification (e.g., 40% or more), as observed in previous studies (e.g., Beale & Keil, 1995; Jacques & Rossion, 2006; McKone, Martini, & Nakaya, 2001), is specific to the 2AFC classification task, which involves two competing face identities with the suppression of non-dominant identities.

With one-third of identity information available, the familiar group recognized more identities than the unfamiliar group. This result suggests that the morphs and their parent faces are perceptually more similar (or closer in face space) for the familiar group than for the unfamiliar group, as the same morph stimuli activate more parent identities for the familiar than unfamiliar group. One possibility is that familiarity with faces enlarges its “attractor field” in face space, so stimuli that matches one-third of its identity information are more likely to fall inside its “attractor field”, leading to the recognition of more faces in a morph. Note that for unfamiliar faces, whether a stimuli fall in/outside the “attractor field” is primarily determined by image-to-image matching, whereas for familiar faces, it can be determined by both image-to-image and image-to-memory matching (cf. Hancock et al, 2000; Johnston & Edmonds, 2009; Young & Burton, 2018).

If one-third of identity information suffices for recognition, what is the minimum identity information that can still preserve its identification? In Experiments 3 and 4, we systematically reduced the contribution of individual parent faces to a morph to assess the minimum information necessary for their recognition. Similar to Experiments 1 and 2, we first tested the limits of familiar face recognition with our colleagues in Experiment 3, and we then tested its generalizability to unfamiliar face recognition with more stimuli, trials, and morph levels in Experiment 4.

**4 Experiment 3**

In Experiment 3 we tested at what point a familiar face is no longer recognizable when its contribution to a morph is gradually reduced. To this end, we created face morphs using two to ten parent faces so that their individual contributions dropped from 50% to 10%. We then assessed how well people can identify individual parent faces at each level of contribution. In contrast to the “whole report” task used in Experiments 1 and 2, where we asked participants to identify all parent faces in one go, Experiments 3 and 4 used a “partial report” task, where in each trial participants assessed whether a single face was a parent of the morph or not. Specifically, we showed participants a morph together with a test face and asked them to decide whether the test face was used to create the face morph (which was the case for half of the trials). If recognition of individual identities requires less than one-third of identity information, participants might show above-chance performance for face morphs created with more than three identities. Such ability was expected to reach its limits when the contribution of individual identities to a morph was reduced to 10% (e.g., 11%, Leopold et al., 2001). Similar to Experiment 1, Experiment 3 used our colleagues’ faces to create the stimuli and tested colleagues who were personally familiar with these faces, which allowed us to test how much identity information is minimally required to recognize familiar faces.

## 4.1 Method

### 4.1.1 Participants

Thirty colleagues from the Max-Planck Institute for Biological Cybernetics took part in this study (11 females; age was not asked individually but ranged roughly between 25 to 60 years old). All were familiar with the identities of the target faces used in the study. Participation was voluntary and no reimbursement was offered.

Power analysis indicated that for a one-way repeated measure design (8 measurements, *f* = 0.25, α =.05, *r*-repeated measures = 0.50), 16 participants are needed to achieve a statistical power (1-*β*) of 0.80, and 23 participants to reach a statistical power of 0.95. Note that among the 30 participants we tested, 15 of them had participated in Experiment 1 and three in Experiment 2. As Experiments 1 and 3 differed in terms of stimuli and tasks and were conducted about one year apart, it is less likely that their participation in Experiment 1 would affect their performance in Experiment 3.

### 4.1.2 Stimuli

Based on the same set of 20 parent faces used in Experiment 1, we create face morphs using 2, 3, 4, 5, 6, 7, 8 or 10 faces (hereafter referred as ***morph level***). All morphs were turned slightly to the side (20°) so that important facial information about the shape of the jaw, cheeks and nose was available. In a morph, all parent faces contributed equally, with contributions ranging from 50% at morph level 2 to 10% at morph level 10. All morphs and test faces were shown on a grey background (512 by 512 pixels, with each face subtending about 9°x14°). For each of the eight morph levels, we created eight morphs (four male and four female morphs). During this experiment, we showed each face morph twice, once in a *parent trial*, in which a face morph is displayed togetherwith one of the parent faces used to create the morph, and once in a *non-parent trial*, in which a face morph is displayed together with a same-sex face that was not used to create that morph (**Figure 5**). In addition to the 20 parent identities used to create the morphs, we used the faces of an additional four male and four female colleagues as test faces for some of the non-parent trials (e.g., at morph level 10).

A collage of different facial expressions

Description automatically generated

**Figure 5**. Examples of trials for three of the eight morph levels (ML) used in Experiment 3. The left panel shows *parent* trials, including a face morph (left side) and a parent face (right side) used to create the face morph. The right panel shows *non-parent* trials, showing a face morph (left side) and a test face (right side) that was not used to create the face morph. Face morphs used for the three morph levels, ML2, ML4 and ML8, were created with 2, 4, and 8 faces, respectively, with each face contributing 50%, 25%, and 12.5% of information to the morph respectively.

### 4.1.3 Procedure

For each trial, the participants’ task was to decide whether the face shown on the right (i.e., test face) was one of the parent faces used to create the face morph shown on the left or not. There were 128 trials in total [2 (male vs. female faces) × 2 (parent vs. non-parent trials) × 8 (morph levels) × 4 trials per condition]. As in Experiment 1, to accustom participants to the appearance of their colleagues’ faces used to create the face morphs, once they had been informed about the task, we showed them the scanned faces of all 14 female and 14 male colleagues used as test faces. These faces were shown individually on the screen for 3 s. Participants then completed two blocks of 64 trials each. The first block displayed only female morphs; the second one showed only male morphs.

Similar to Experiment 1, due to colleagues’ availability, most of them (28 of 30) took part in the study in a group with face morphs presented on a projection screen (3 to 5 m from participants seating position). Another two participants performed the same task individually in front of a computer monitor. Participants entered their responses on an answer sheet, where each trial number (i.e., 1-128, as shown on the screen) was followed by an answer box that they could tick “Yes” or “No” as a response. Each trial was displayed for 5 s followed by a blank screen until a response was made. At the end, participants were given another sheet showing all parent faces and were asked to indicate the faces that were unfamiliar to them.

## 4.2 Results

Data of trials that showed a parent face that a participant indicated as unfamiliar (217 of 3840 trials, 5.7%) or recorded no response (9 trials, 0.2%) were excluded from the analysis. **Figure 6** shows participants’ performance measured as response sensitivity (*d’,* e.g., Dunn, 2010), hit rate, false alarm rate, and response criterion (i.e., bias) as a function of available identity information (see **Figure A1** in **Appendix C** for the relationship between performance and morph level). One-way repeated measures ANOVAs revealed strong effects of morph levels on *d*’, hit rate and bias [*F*(7,203) ≥ 27.96, *p*<.001, *ηp*2 ≥ .491] with significant linear trends [*F*(1,29) ≥ 205.75, *p* <.001, *ηp*2 ≥ .876]. For false alarm rate, the effect of morph level was significant (*F*(7,203) = 2.85, *p*=.007, *ηp*2 =.089), but no significant linear trend was observed (*F* < 1).

Response sensitivity (*d*’) decreased with decreasing face identity information in the morph, showing a linear relationship (*R2*= .93; *RMSE* = 0.20, *F*(1,6) = 79.05, *p* < .001; **Figure 6A**; see **Table A3** in **Appendix D** for results of nonlinear fitting of d’, hit and response bias). One-sample t-tests revealed that performance was significantly above-chance (i.e., *d*’ = 0) up to morph level 8, all *t*(29) > 4.33, *p* < .001, Cohen’s *d* > 0.79; but not at morph level 10, *t*(29) = 1.98, *p* = .057, Cohen’s *d* = 0.36. This result indicates that people still show above-chance recognition with only 12.5% of face identity information (i.e., at morph level 8) but fail to do so with 10% of identity information (i.e., performance is equivalent to guessing at morph level 10).

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

**Figure 6.** Performance measured as response sensitivity, hit and false alarm rate, and response bias in Experiment 3. Lines represent best fits. Error bars represent SEM. For all fitted equations, X refers to the proportion of identity information (e.g., 0.1 for 10% shown in the x-axis).

The decrease of response sensitivity with decreasing identity information was driven by a linear decrease of hit rate (*R2*= .93; *RMSE* = 0.06, *F*(1,6) = 76.48, *p* < .001; **Figure 6B**). A 10% reduction in identity information led to a decrease of about 0.5 in response sensitivity and a 0.14 drop for hit rate. In contrast, false alarm rate remained largely unchanged across our manipulation of identity information (*R*² = 0.15; *RMSE* = 0.03, *F*(1,6) = 1.06, *p* = .343; **Figure 6C**). Finally, participants became increasingly more conservative (i.e., tendency to make “No” responses) with decreasing identity information (*R2*= .90; *RMSE* = 0.10, *F*(1,6) = 54.36, *p* < .001; **Figure 6D**). One-sample t-tests revealed a significant liberal bias (tendency to make “Yes” responses) with 50% of identity information (i.e., at morph level 2), *t*(29) = 2.89, *p* = .004, and significant conservative bias when identity information reduced to 25% or less (i.e., morph levels 4 and higher), all *t*(29) ≥ 2.61, *p* ≤ .007. Significant response biases that observed with 20-50% of identity information (i.e., morph levels 2 to 5) did not survive Bonferroni correction.

Previous studies found that kinship decisions (e.g., determining whether two faces are parent and child or are siblings) can be predicted by perceptual similarity between the test faces (DeBruine et al., 2009; Maloney & Dal Martello, 2006). To examine if the observed decreasing response sensitivity (and hit rate) is similarly mediated by the similarity between the test faces, we computed image-based similarity between the morph and the test face for every trial and examined its relationship with morph levels and participants’ performance. Image-based face similarity was measured as Gabor dissimilarity (Lades et al., 1993), which has been shown to correlate with perceptual similarity between faces (Yue et al., 2012; see also Bülthoff & Zhao, 2020, 2021; Dobs, Bülthoff, & Schultz, 2016; for details of calculation of Gabor similarity).

Mean Gabor dissimilarity for parent and non-parent trials at each morph level are shown in **Figure 7A**. A fixed effect linear mixed model (LMM) analysis (“Gabor dissimilarity ~ 1 + Trial type\*Morph level”) revealed a main effect of trial condition, *F*(1, 112) = 44.19, *p* < .001, indicating that the two faces shown in the non-parent trials (223.44±3.25) differ more from each other than those in the parent trials (192.85±3.25). The main effect of morph level was not significant, *F*(7, 112) = 1.72, *p* = .112, but the interaction between morph level and trial type was significant, *F*(7, 112) = 2.75, *p* = .011. Higher morph levels led to higher image differences between the two faces in the parent trials, but not in the non-parent trials. When comparing image similarity between parent and non-parent trials (independent sample *t*-tests), significant differences were found for morph levels 2-4 (though morph level 4 did not survive the Bonferroni correction), all *t*(14) ≥ 2.33, *p* ≤ .035, Cohen’s *d* ≥ 1.17, but not for higher morph levels 5-10, all *t*(14) ≤ 1.91, *p* ≥ .077, Cohen’s *d* ≤ 0.95. These results indicate that the computational resemblance between the morph and its parent faces diminished gradually with increasing morph levels, down to being equivalent to the resemblance between a morph and a non-parent test face.

A graph of different stages of a parent

Description automatically generated with medium confidence

**Figure 7**. Computed image similarity for Experiment 3. (A) Gabor dissimilarity between faces in a trial as a function of trial type and morph level. (B) Differences between image similarity in parent and non-parent trials correlate with response sensitivity at individual morph levels. (C and D) Relationship between image similarity and hit and false alarm rates. Error bars are SEMs and trend lines are best fits of linear regression.

To investigate whether image similarity at individual morph levels underlies participants’ performance, we conducted two correlation analyses. First, we examined whether participants’ response sensitivity reflects the difference between face similarity in the parent trials (i.e., signal-present) and the non-parent trials (i.e., signal absent). Second, we tested whether hit and false alarm rates (i.e., proportion of yes responses) were determined by face similarity at each morph level. As shown in **Figure 7B**, mean differences in face similarity between parent and non-parent conditions were significantly correlated with the mean *d*’ observed at individual morph levels, *r* =.953, *p* <.001. This result suggests that participants’ responses were mediated by face similarity; larger differences between face similarity in the parent and non-parent conditions led to higher response sensitivity.

We then performed separate correlation analyses on the responses to the parent and non-parent trials. For parent trials, mean hit rates at individual morph levels were significantly correlated with mean Gabor dissimilarity between the faces in each trial (**Figure 7C**), *r* = -.952, *p* < .001. In contrast, for the non-parent trials, false alarm rate remained constant regardless of face similarity (**Figure 7D**), *r* = .065, *p* = .878. These results indicate that the tendency to respond “Yes” gradually diminished with decreasing similarity between the faces in a trial, and it appears to stop decreasing when it reached a threshold (e.g., when Gabor dissimilarity between two faces is around 205, **Figure 7C and 7D**).

## 4.3 Discussion

Experiment 3 shows two main findings. First, the ability to recognize a parent face rapidly decreases when its contribution to a face morph is reduced. Moreover, such ability appears to reach its limit at 10% of identity information (i.e., morph level 10), where performance is at chance level. This finding echoes Leopold et al (2001) that people need about 11% of identity information for face identification. Second, reduced stimulus-based similarity at higher morph levels can account for the decreasing ability to identify a parent face from a morph. Participants adopted a more conservative criterion (i.e., less tendency to respond “Yes”) when the physical similarity between two faces in a trial was gradually reduced. This is mirrored by a stable false alarm rate and a decreasing hit rate with increasing morph levels. These results indicate that idiosyncratic information of parent faces in a morph is diluted to a non-recognizable level when ten identities are mixed in a morph. In Experiment 4, we tested whether these findings generalize to the recognition of unfamiliar faces.

**5 Experiment 4**

Experiment 4 tested how much identity information is required to identify an unfamiliar identity in a morph. In contrast to Experiment 3, where stimuli, participants, and the number of trials we could test were limited by the number and availability of our colleagues, in Experiment 4, we systematically tested how well people can recognize an unfamiliar identity from a morph using the same task but with more stimuli, trials, and morph levels. By comparing the results obtained here with those of Experiment 3, we were able to further examine whether familiarity with faces affects our ability to recognize multiple faces in a morph.

## 5.1 Method

### 5.1.1 Participants

Twenty-four participants (13 females; mean age: 27.6 years, range = 19-55) took part in this experiment. None of them were familiar with the faces used to create the stimuli. Written informed consent was obtained before the experiment. Participants received small monetary payment for their participation. Power analysis indicated that for a one-way repeated measures design (10 measurements, *f* = 0.25, α =.05, *r*-repeated measures = 0.50), 14 participants are needed to achieve a statistical power (1-*β*) of 0.80, and 20 participants to reach a statistical power of 0.95.

### 5.1.2 Stimuli

Stimuli were created using 160 female faces from the Max Planck Face database. We created 40 morphs as described in Experiment 3 at each of 10 morphing levels (i.e., morphs created using 2, 3, 4, 5, 6, 7, 8, 10, 12, or 16 parent identities), resulting in 400 morphs in total. We showed each morph twice during the study, once in the parent trial and once in the non-parent trial, yielding 800 trials. To control for the potential influence of age differences between faces in each trial, we grouped the 160 faces so that all parent faces used to create a specific face morph had similar ages, and both the morph and the test faces shown in a trial also shared a similar age.

### 5.1.3 Procedure

The study was controlled using the E-prime software. Participants performed the same task as in Experiment 3. After having been instructed about the task, participants had 20 practice trials to familiarize themselves with the task before completing the 800 experimental trials presented in a random order. There were self-timed breaks every 100 trials. Each trial began with a fixation screen (200 ms), followed by a pair of faces showing the morph on the left and the test face on the right. This face pair was displayed for a maximum of 5 s, then a blank screen followed and remained displayed until a response was made. Participants responded by pressing keys corresponding to “Yes” or “No” on a keyboard and they could respond as soon as the trial started. The next trial followed immediately after a key press, so two successive trials were separated either by a short fixation screen (when a response was made during the 5 s display time) or by a blank screen and a fixation screen (when a response was entered after the 5 s had passed).

## 5.2 Results

The overall responses showed a pattern similar to that in Experiment 3 (**Figure 8**,see **Figure A2** in **Appendix C** for the relationship between performance and morph level). One-way repeated measures ANOVAs revealed strong effects of morph levels on *d*’, hit rate and bias, [all *F*(Greenhouse-Geisser corrected) *≥* 40.20, *p* <.001, *ηp*2 ≥ .636] with significant linear trends [*F*(1,23) ≥ 76.03, *p* <.001, *ηp*2 ≥ .768]. For false alarm rate, the effect of morph level was significant [*F*(9,207) = 5.19, *p* <.001, *ηp*2 =.184], but no significant linear trend was observed [*F*(1,23) = 1.39, p =.25, *ηp*2 = .057].

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

**Figure 8**. Performance measured as response sensitivity, hit and false alarm rate, and response bias at each morph level for Experiment 4. Error bars represent SEM. For all fitted equations, X refers to the proportion of identity information (e.g., 0.1 for 10% shown in the x-axis).

Response sensitivity (*d*’) decreased with decreasing identity information (*R2* =.96, *RMSE*= 0.12; *F*(1,8) = 175.93, *p* < .001; **Figure 8A**). One-sample *t*-tests showed above chance performance for all levels of identity information tested [6.25-50%, all *t*(23) ≥ 3.67, *p* < .001, Cohen’s *d* ≥ 0.75] except the 8.33% one at morph level 12, *t*(23) = 0.37, *p* = .715, Cohen’s *d* = 0.08. Therefore, participants still showed above-chance recognition with 10% of identity information, but their performance dropped to guessing level when the contribution of identity information was below 10% (e.g., at morph level 12).

Reduction in identity information (with increasing morph level) affected hit and false alarm rates differently (**Figure 8B and 8C**). Hit rates decreased linearly with decreasing identity information (*R2* =.96, *RMSE*= 0.04, *F*(1,8) = 174.85, *p* < .001). A 10% reduction in identity information led to about a 0.4 decrease in response sensitivity and a 0.12 drop for hit rate. In contrast, false alarm rates remained relatively constant across different levels of identity information (*R²* = .002, *RMSE* = 0.03, *F*(1,8) = 0.02, *p* = .896). For response criterion (**Figure 8D**), participants’ responses were increasingly more conservative with decreasing identity information (i.e., increasing morph level, *R2*= .92, *RMSE*= 0.07, *F*(1,8) = 87.83, *p* < .001). One-sample t-tests revealed a significant liberal bias (i.e., tendency to say “Yes”) with 50% of identity information [i.e., morph level 2, *t*(23) = 2.64, *p* = .015, Cohen’s *d* = 0.54] and significant conservative bias (i.e., tendency to respond “No”) when identity information dropped to 14% or less [i.e., morph levels 7 to 16, all *t*(23) > 2.70, *p* ≤ .013, Cohen’s *d* ≥ 0.55]. Only significant biases observed with 12.5% or less of identity information (i.e., morph levels 8 to 16) survived the Bonferroni correction.

We then performed the same analysis using stimulus-based similarity as in Experiment 3 (**Figure 9**). Image dissimilarity between two faces in a trial was higher for non-parent (220.78±1.90) than for parent trials (194.06±1.73), but the difference between both conditions diminished gradually with increasing morph levels (**Figure 9A)**. This was supported by a LMM fixed effect analysis (“Gabor dissimilarity ~ 1 + Trial type\*Morph level”), which revealed a main effect of trial condition, *F*(1, 780) = 108.36, *p* < .001, and a significant interaction between morph level and trial condition, *F*(9, 780) = 7.11, *p* < .001. The main effect of morph level was not significant, *F*(9, 780) = 1.60, *p* = .110. Significant differences between parent and non-parent trials were observed at all morph levels, all *t*(78) ≥ 2.15, *p* ≤ .035, Cohen’s *d* > 0.48, except for morph levels 10 and 12, both *t*(78) ≤ 0.920, *p* ≥ .360, Cohen’s *d* < 0.21. Only the differences at morph levels 2-5 survived Bonferroni correction. Note that significant image-based difference at morph level 16 may account for the surprising finding that performance was still above chance with less than 10% of identity information available (**Figure 8A**).

A diagram of a number of different types of data

Description automatically generated with medium confidence

**Figure 9**. Computed image similarity for Experiment 4. (A) Gabor dissimilarity between faces in a trial as a function of trial condition and morph level. (B) Differences between image similarity in parent and non-parent trials correlate with response sensitivity at individual morph levels. (C and D) Relationship between image similarity and hit and false alarm rates. Error bars are SEMs and trend lines are best fits of linear regression.

Response sensitivity (*d*’) was significantly correlated with the difference in face similarity between parent trials (i.e., signal-present) and non-parent trials (i.e., signal absent), *r* = .97, *p* <.001 (**Figure 9B**). Larger differences in face similarity between parent and non-parent conditions led to higher response sensitivity. Hit rates correlated with the similarity between two faces in a parent trial, *r* = -.926, *p* < .001 (**Figure 9C**), however, face similarity in the non-parent trials showed no such effect on false alarm rate, *r* = -.148, *p* = .684 (**Figure 9D**). These results indicate that people become more reluctant to respond “Yes” when two faces in a trial become more different, and once such a difference reaches a threshold (e.g., around 205 in Gabor dissimilarity, **Figure 9C** and **9D**), the proportion of “Yes” responses remain relatively constant.

Finally, to test whether familiarity with faces affects participants' performance, we combined data from Experiments 3 and 4 and performed a LMM fixed effect analysis with morph level and Experiment as fixed factors (“Responses ~ 1 + Experiment\*Morph level”). Since morph levels 12 and 16 were only tested in Experiment 4, data from these two morph levels were excluded from this analysis. We found significant main effects of participant group on response sensitivity (*d*’), *F*(1,416) = 20.71, *p* < .001, false alarms, *F*(1,416) = 39.57, *p* < .001; response bias, *F*(1,416) = 12.61, *p* < .001; but not on hit rates, *F*(1,416) = 1.00, *p* = .430. Overall, the familiar group showed higher response sensitivity (1.04±0.04 vs 0.80±0.04), lower false alarms (0.22±0.01 vs 0.32±0.01), and higher conservative bias (0.35±0.03 vs 0.17±0.04) than the unfamiliar group (see **Figure A3** in **Appendix E** for details).

Significant influence of morph level was observed on *d*’, *F*(7,416) = 62.86, *p* < .001, hit rate, *F*(7,416) = 41.56, *p* < .001, and response bias, *F*(7,416) = 15.05, *p* < .001, but not on false alarms, *F*(7,416) = 1.28, *p* = .258. The interaction between morph level and participants group was only significant for *d*’, *F*(7,416) = 2.78, *p* = .008; no interaction was found for hit rate, *F*(7,416) = 1.00, *p* = .430; false alarm, *F*(7,416) = 1.09, *p* = .369; and bias, *F*(7,416) = 0.55, *p* = .800. Follow-up independent t-test on *d*’ revealed that the familiarity advantage was significant at morph levels 2, 3, 6 and 7, all *t*(52) ≥ 2.38, *p* ≤ .021, Cohen’s *d* > 0.65 (though difference at morph level 7 did not survive the Bonferroni correction). These results indicate that familiarity enhances participants’ ability to recognize multiple identities in a face morph, but such enhancement is prominent at lower morph levels and vanishes at high morph levels.

## 5.3 Discussion

Experiment 4 shows three main findings. First, recognizing unfamiliar faces from a morph follows the same principle as recognizing familiar ones. Gradual reduction of parent identity information in a morph leads to increased difficulty in recognizing them. Second, increased image differences between a face morph and a parent face underlie the observed decreased recognition performance with decreasing identity information. Third, familiarity with faces facilitates their recognition in a morph, but this familiarity effect is more pronounced when 16% or more information is available about a parent identity (i.e., when a morph is created with up to 6 faces).

A seemingly surprising finding is that our unfamiliar group showed above-chance performance at morph level 16, with 6.25% contribution from individual parent identities. Note that our familiar group lost their ability to do so with 10% of identity information (Experiment 3). A detailed examination of image-based similarity at high morph levels revealed that this surprising result is due to unexpected high image differences between parent and non-parent trials. As summarized in Table 1, across Experiments 3 and 4, participants were sensitive to such image-based differences and an above-chance level decision could be consistently made with as small as an 8% difference (about 20 units in terms Gabor dissimilarity) between parent and non-parent trials. At morph level 16, image difference between parent and non-parent trials (8% or 19 units) was nearly identical to that at morph level 8 (8.1% or 20 units, see also **Figure 9A**); both values are above the threshold for above-chance performance.

**Table 1**. Image-based similarity and performance around the threshold.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Morph level | | Gabor similarity: Non-parent trials | Gabor similarity: Parent trials | Difference (%) | Mean d’ (Hit/FA) |
| Exp.3 | 8 | 229 | 207 | 22/9.6% | \*0.44 (0.37/0.23) |
| 10 | 223 | 213 | 10/4.5% | 0.19 (0.28/0.21) |
| Exp.4 | 8 | 225 | 205 | 20/8.1% | \*0.46 (0.43/0.28) |
| 10 | 215 | 208 | 7/3.0% | \*0.31 (0.41/0.32) |
| 12 | 204 | 208 | -4/2.0% | -0.03 (0.34/0.33) |
| 16 | 222 | 204 | 19/8.0% | \*0.24 (0.36/0.28) |

*Notes. \*, denotes response sensitivity (d’) that is significantly above chance level. Differences are given in Gabor dissimilarity units and as percentages relative to the non-parent trial condition*.

**6 General Discussion**

The primary goal of the present study was to investigate how much identity information in a face morph is required for face recognition. By allowing people to report multiple recognized face identities from a morph created with three faces (Experiments 1 and 2), our participants recognized half of them based on their memory of these faces, and they recognized two out of three faces when perceptual information about these faces is also available. By systematically reducing the contribution of individual identities to a facial morph (e.g., from 50% to 6.25%, Experiments 3 and 4), participants could consistently recognize a face in a morph at an above-chance level even when the face contributes only 12.5% of information to that morph. The ability to identify a face in a morph reaches its limit when its identity information is further reduced (e.g., 10% or less). Consistent with Leopold et al. (2001), these results indicate that better-than-chance face identification can be reliably achieved with only a small fraction of identity information.

Our results are different from previous studies that employed a 2AFC face classification task, where participants had to decide which one of two parent faces was displayed in a morph (Beale & Keil, 1995; Campanella et al., 2003; Chauhan et al., 2020; Levin & Beale, 2000; McKone et al., 2001). These studies have consistently demonstrated that people require 40% or more of identity information to recognize a face (i.e., otherwise, a target face is reported less than 20% of the time). When only 20% of parent identity information is available, people rarely recognize the parent face in a morph (i.e., only reported it less than 10% of the time). Our study suggests that this higher proportion of identity information required for recognition is specific to the 2AFC task, as people have to report only the perceptually dominant face even though they may simultaneously recognize the other competing identity. Therefore, while the 2AFC classification task is ideal for determining the categorical boundary between two competing face identities, it is not a sensitive tool to measure the minimal identical information required for face recognition.

The minimal identity information required for face recognition (i.e., deciding if a face displays some idiosyncratic information of a known person) is similar to that required for face discrimination (i.e., deciding if two faces are different). When different amounts of information from two faces are morphed together, people can consistently discriminate between two morphs that differ by 20% of information (e.g., 20% face A and 80% face B versus 40% face A and 60% face B), whether these morphs are created with familiar or newly learned faces (showing a 55% or higher accuracy; Bülthoff & Newell, 2004; Campanella et al., 2003; Kikutani et al., 2008, 2010; Beale et al., 1995; Levin & Beale, 2000). When the differences between face morphs were manipulated at a more fine-grained level, people needed a difference of 6-14% between two faces to discriminate between them (Bülthoff & Zhao, 2021; Gao & Wilson, 2013; McKone et al., 2001; Wilson et al., 2002). Here we show that participants are no longer able to consistently identify a face when less than 10% of its identity information is displayed, and that they need an 8% image-based difference between two faces to determine correctly whether two faces share some identity information (i.e., Table 1). These results suggest that the minimum identity information required for face recognition is close to the threshold used to discriminate between two different faces (e.g., 6-8%, Gao & Wilson, 2013; Wilson et al., 2002; 7-14%, McKone et al., 2001).

While morphing has been frequently used to vary the amount of identity information in an image, perceptual similarity between a face morph and a parent face may not always align with their physical similarity (e.g., Busey, 1998; Bülthoff & Zhao, 2020; 2021; Tanaka et al., 1998). For instance, Tanaka and colleagues (1998) showed that when a typical face and a distinctive face are morphed equally together, people perceive the morph as more similar to the distinctive face than the typical face. In the present study, we did not measure the perceptual distinctiveness of individual faces. It is possible that our face morphs, particularly those created with a small number of faces, may bear a stronger resemblance to some parent faces than to others depending on their distinctiveness. If the shape and textural information our male faces are more distinctive than that of female faces, this may explain why our participants showed better performance for male morphs than for female or mixed-gender morphs (Experiments 1 and 2). For Experiments 3 and 4, as the same parent faces were used for multiple morph levels and multiple parent faces were used for the same morph level, the potential influence of face distinctiveness could be cancelled out (e.g., in terms of the effects of familiarity or morph level) or accounted for by our image-based similarity analysis (e.g., when parent and non-parent trials are not distinguishable). Further investigation is needed to examine whether the minimum identity information required for recognition similarly applies to typical and distinctive faces.

Although the present study does not test kinship recognition, it shows some parallel findings to those found in previous kinship studies. For instance, children’s faces are perceived as more similar to a parent than to a non-parent face by unfamiliar observers (Alvergne, Faurie, & Raymond, 2007; Bressan & Dal Martello, 2002; Christenfeld & Hill, 1995; McLain et al., 2000; Oda, Matsumoto-Oda, & Kurashima, 2005). Moreover, kinship detection based on faces by unfamiliar observers is better for faces of close (e.g., siblings) than distant kinship (e.g., cousins). For instance, Kaminski et al. (2009) showed that the more two faces were closely related, the more they were assessed as kin by people who were unfamiliar with them. Above-chance detection of kinship can be made for faces of siblings (sharing 50% of genes), grandparents and grandchildren (25%), but not for faces of aunt/uncle and nephews (25%) or cousins (12.5%). Even in the last two conditions, these faces were still assessed as more related than two faces with no kinship.

Like the above real-life kinship studies, we showed that faces in the parent trials (i.e., analogous to faces with kinship) are computationally more similar to each other than faces in the non-parent trials (i.e., analogous to faces without kinship). Moreover, when two faces share more identity information (e.g., 50% for faces in parent trials at morph level 2; analogous to a “close kinship”), they are computationally more alike than when two faces share less identity information (e.g., 12.5% for faces in parent trials at morph level 8; analogous to a “distant kinship”). While the results of Kaminski et al (2009) suggest that a minimum of a quarter of shared genes (e.g., between grandparents and grandchildren) is required to detect real-life kinship, our results suggest that people can detect shared identity between two faces with one-eighth of identity information available. Note that real-life kinship studies often used highly heterogeneous test faces (especially in terms of age, Bressan & Dal Martello, 2002; DeBruine et al., 2009; Kaminski et al., 2009; Maloney & Dal Martello, 2006; McLain et al., 2000), which contrasts with the high homogeneity of our stimuli. Although the processes involved in our tasks and kin recognition are likely to be different, these parallel findings indicate that people are sensitive to the idiosyncratic resemblance between two faces, whether such resemblance is due to biological kinship or face morphing (cf. Christenfeld & Hill, 1995).

Consistent with the well-documented differences between the recognition of familiar and unfamiliar faces (Hancock et al., 2000; Johnston & Edmonds, 2009; Megreya & Burton, 2006; Young & Burton, 2018), we found that familiarity with faces modulates the minimum identity information required for face recognition. Familiarity enhances recognition of multiple face identities mixed in one image (Experiments 1 and 2) and improves the detection of parent identities in a morph (e.g., for morphs created with 2 to 6 faces; Experiments 3 and 4). Compared to a newly formed face representation in face space (i.e., for the unfamiliar group), where its initial “attractor field” is so limited that a different image of the same face could fall outside (e.g., Hancock et al., 2000; Young & Burton, 2018), representation of a familiar face appears to have a larger “attractor field” that tolerates variations in image properties and reductions in idiosyncratic information (e.g., Bindermann et al., 2013; Gilad-Gutnick et al., 2012; Lander et al., 2001). When face morphs contain only a small proportion of parent identity information, differences in the range of their “attractor field” may make familiar identities in a face morph more likely to be detected and recognized than unfamiliar ones. Note that this familiarity advantage diminishes when more than six faces are mixed in a morph.

Our speculation that personally familiar faces have a larger “attractor field” echoes a recent finding that personal familiarity warps face representation in face space (Chauhan et al., 2020). When participants are forced to report only one identity from a two-identity morph (50% from each face), they adopt a more conservative criterion for familiar than unfamiliar faces, suggesting that changes in the appearance of familiar faces are amplified in perceptual space compared to that for unfamiliar faces (Chauhan et al., 2020; see also Beale & Keil, 2000). Our participants also showed stronger conservative response criteria for familiar than unfamiliar faces. Moreover, across Experiments 3 and 4, we observed equivalent hit rates together with consistent lower false alarm rates for familiar than unfamiliar faces. These findings suggest that familiarity with faces amplifies their perceptual distances to neighboring faces by pushing them apart in face space (**Appendix E**, **Figure S3**).

If familiarity with a face enlarges its “attractor field” or amplifies its representational space (e.g., Chauhan et al., 2020), quantifying such changes may offer a new perspective on why familiar and unfamiliar faces are processed differently. Differences between processing familiar and unfamiliar faces have often been viewed as qualitative (e.g., Hancock et al., 2000; Megreya & Burton, 2006; Young & Burton, 2018). For instance, it has been proposed that representations of familiar faces are “robust” or “abstract” (e.g., Burton et al., 2005; Tong & Nakayama, 1999; Ramon & Gobbini, 2018), whereas representations of unfamiliar faces are image-based or object-like (e.g., Megreya & Burton, 2006). What quantitative changes may underlie such qualitative differences remains to be elucidated. An increased “attractor field” for the representations of familiar faces may offer one such quantitative dimension, which could help explain (i) why familiar faces can be recognized with much more coarse facial information than unfamiliar faces (Ramon et al., 2015; Watier & Collin, 2009; see also Bindermann et al., 2013; Lander et al., 2001), (ii) why familiarity with a person leads to a higher rating of likeness to a variety of their facial images (Ritchie, Kramer & Burton, 2018), and (iii) why familiarity with someone helps identification of their relatives or other faces similar to them (Hancock, 2021). Note that familiarity with faces changes both perceptual and social information associated with their representations (Schwartz & Yovel, 2019; Ramon & Gobbini, 2017), and the potential change of “attractor field” reflects the tuning of perceptual, but not social, information.

Finally, our study suggests that analyzing stimulus-based physical similarity may help quantify information required for face identification in an objective and sensitive manner. Different methods have been used to manipulate the amount of face identity information in an image, either directly (e.g., morphing between identities and morphing along the identity trajectory) or indirectly (e.g., blurring, pixelating, or scrambling facial parts). For instance, it has been shown that people are remarkably good at recognizing faces from dramatically pixelated or blurred images (e.g., 10-20 pixels/face; Bindermann et al., 2013; Lander et al., 2001; 8-16 cycles per face; Costen et al., 1996; Ramon et al., 2015; or 3.5 cycles per face for discrimination, Gilad-Gutnick et al., 2012). While the magnitude of these image degradations can be calculated as information loss, how such information loss relates to the amount of reduction in identity information remains unclear. Future research may use physical and perceptual similarity to quantify identity information retained within these highly pixelated or blurred face images.

In summary, the present study demonstrates that people can recognize multiple face identities in a morphed image and can consistently recognize a face even when only 12.5% of its identity information is displayed in an image. Therefore, people are able to recognize at least eight identities that are mixed equally in one face image. Furthermore, familiarity with faces enhances their recognition in a morph, indicating that representation of familiar faces tolerates more reduction of identity information than that for unfamiliar faces. These results not only help quantify the minimum information required for face recognition but also add new evidence to the well-established representational differences between familiar and unfamiliar faces.

**CRediT authorship contribution statement**

**Mintao Zhao** and **Isabelle Bülthoff** both contributed to the Conceptualization, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, and Project administration of this research.

**Supplementary material and data availability**

Raw data can be found in the supplementary materials associated with this article. All data and the related analysis files are also publicly available at https://osf.io/kc6xz/.

**Appendices**

**Appendix A. Chance-level performance for Experiment 1 and 2.**

The probability of selecting a parent face when choosing 3 faces from an array of 20 test faces that contains 3 parent faces follows the hypergeometric distribution.

The expected number of correctly selected parent faces by chance can be calculated using the expected value formula for the hypergeometric distribution:

Where *n* is number of faces chosen (i.e., 3), *K* is the number of parent faces in the array (i.e., 3), and *N* is the total number of test faces (i.e., 20).

So, by chance, participants are expected to accurately select 0.45 parent faces:

3 × (3/20) = 0.45

**Appendix B. Results of Bayesian ANOVAs for Experiment 2**

The Bayesian ANOVAs were performed using the JASP software (https://jasp-stats.org/, van den Bergh et al., 2020). The results showed decisive evidence that (1) familiarity had an effect on recognition performance (BFInclusion = 316.895, Table A1) and (2) differences between Experiments 1 and 2 had an effect on recognition of familiar faces (BFInclusion = 525.297, Table A2).

**Table A1**. Analysis of effects for the 2 (Familiarity) × 3 (Morph Type) Bayesian ANOVA performed on the data obtained in Experiment 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Effects | P(incl) | P(excl) | P(incl|data) | P(excl|data) | BFincl |
| Morph Type | 0.60 | 0.40 | 1.000 | 4.171×10-6 | 159850.343 |
| Familiarity | 0.60 | 0.40 | 0.998 | 0.002 | 316.895 |
| Morph Type × Familiarity | 0.20 | 0.80 | 0.536 | 0.464 | 4.616 |

*Notes.* Effects refer to predictors of interests; P(incl) and P(excl) show the prior inclusion and exclusion probability respectively; P(incl|data) and P(excl|data) show the posterior inclusion and exclusion probability respectively; BFincl shows the inclusion Bayes factor, which quantifies the change from prior inclusion odds to posterior inclusion odds and can be interpreted as the evidence in the data for including a predictor.

**Table A2**. Analysis of effects for the 2 (Experiments) × 3 (Morph Type) Bayesian ANOVA performed on the combined data from the familiar groups in Experiments 1 and 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Effects | P(incl) | P(excl) | P(incl|data) | P(excl|data) | BFincl |
| Morph Type | 0.60 | 0.40 | 1.000 | 2.427×10-5 | 27467.289 |
| Experiment | 0.60 | 0.40 | 0.999 | 0.001 | 525.297 |
| Morph Type × Experiment | 0.20 | 0.80 | 0.262 | 0.738 | 1.423 |

**Appendix C. Performance as a function of morph level observed in Experiments 3 and 4.**

**Experiment 3**. Response sensitivity decreased nonlinearly with increasing morph level; the decrease was rapid at lower morphing levels and slowed down at higher morph levels, which fitted nicely with a power function (*R2*= .93; *RMSE* = 0.20, Figure A1A). The decrease of response sensitivity with increasing morph level was mainly driven by a decreasing hit rate (Figure A1B), which dropped nonlinearly with increasing morph level and fitted nicely with a power function (*R2*= .96; *RMSE* = 0.04). In contrast, false alarm rate remained largely unchanged across morph levels (*R*² = 0.015; *RMSE* = 0.04, Figure A1C). Finally, participants’ response criterion become increasingly more conservative (i.e., tendency to make “No” responses) with decreasing identity information (i.e., increasing morph level, (*R2*= .97; *RMSE* = 0.06, Figure A1D).

A graph of a number of different levels

Description automatically generated with medium confidence

**Figure A1.** Response sensitivity (*d*’), hit and false alarm rates, and response bias as a function of morph level observed in Experiment 3.

**Experiment 4**. Response sensitivity (*d*’) decreased nonlinearly with increasing morph level, a pattern captured nicely with a power function (*R2* =.95, *RMSE*= 0.12; Figure A2A). Hit rates decreased with increasing morph level, which fits nicely with a power function (*R2* =.98, *RMSE*= 0.03; Figure A2B), whereas false alarm rates remained relatively constant across morph levels (*R²* = .07, *RMSE* = 0.03; Figure A2C). For response criterion (Figure A2D), participants were increasingly more conservative in their responses with increasing morph level (*R2*= .96, *RMSE*= 0.07).

A graph of a function

Description automatically generated with medium confidence

**Figure A2.** Response sensitivity (*d*’), hit and false alarm rates, and response bias as a function of morph level observed in Experiment 4.

**Appendix D. Linear and nonlinear fitting of responses observed in Experiments 3 and 4.**

**Table A3**. Curve fitting results for *d*’, hit rate and bias for Experiments 3 and 4.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Experiment 3* | | | | | | |
|  | Sensitivity (d’) | | Hit rate | | Bias | |
| Function | *R2* | *RMSE* | *R2* | *RMSE* | *R2* | *RMSE* |
| f(x) = ax+b | **0.93** | **0.20** | 0.93 | 0.06 | **0.90** | **0.10** |
| f(x) = a\*exp(-x\*b)+c | 0.93 | 0.21 | **0.97** | **0.04** |  |  |
| f(x) = a\*x^b |  |  |  |  | 0.90 | 0.10 |
| *Experiment 4* | | | | | | |
|  | Sensitivity (d’) | | Hit rate | | Bias | |
| Function | *R2* | *RMSE* | *R2* | *RMSE* | *R2* | *RMSE* |
| f(x) = ax+b | **0.96** | **0.12** | 0.96 | 0.04 | **0.92** | **0.09** |
| f(x) = a\*exp(-x\*b)+c | 0.96 | 0.13 | **0.98** | **0.03** |  |  |
| f(x) = a\*x^b |  |  |  |  | **0.67** | **0.18** |

*Note. Better results of goodness of fit for the linear and nonlinear functions are shown in bold.*

**Appendix E. Comparison between performance observed in Experiments 3 and 4.**

As shown in the figure below, the familiar showed equivalent Hit rate to, but lower false alarm rate than, the unfamiliar group. These results suggest that parent and non-parent faces (i.e., analogous to signal and noises in the detection theory) are perceptually more distant from each other for the familiar group than for the unfamiliar group.

A group of graphs showing different types of facial features

Description automatically generated with medium confidence

**Figure A3.** Response sensitivity (*d*’), hit and false alarm rates, and response bias as a function of morph level and familiarity with parent faces observed in Experiments 3 and 4.

**References**

Alvergne, A., Faurie, C., & Raymond, M. (2007). Differential facial resemblance of young children to their parents: who do children look like more? *Evolution and Human Behavior, 28(2)*, 135–144. https://doi.org/10.1016/j.evolhumbehav.2006.08.008

Bachmann, T. (1991). Identification of spatially quantised tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology, 3(1),* 87-103.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57(3),* 217-239. https://doi.org/10.1016/0010-0277(95)00669-X

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., … Wagenmakers, E. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année Psychologique, 120(1),* 73-96. https://doi.org/10.3917/anpsy1.201.0073

Bindemann, M., Attard, J., Leach, A. and Johnston, R.A. (2013), The effect of image pixelation on unfamiliar-face matching. *Applied Cognitive Psychology, 27*, 707-717. https://doi.org/10.1002/acp.2970

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of SIGGRAPH 99, Computer Graphics Proceedings, Annual Conference Series*, edited by Alyn Rockwood (pp. 187–194). Reading, MA: Addison Wesley Longman. https://doi.org/10.1145/311535.311556

Bressan, P., & Dal Martello, M. F. (2002). Talis pater, talis filius: Perceived resemblance and the belief in genetic relatedness. *Psychological Science, 13(3),* 213–218. https://doi.org/10.1111/1467-9280.00440

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognizing the sex and identities of faces. *Perception, 23(7)*, 803–822. https://doi.org/10.1068/p230803

Bülthoff, I., Manno, L., & Zhao, M. (2023). Varying sex and identity of faces affects face categorization differently in humans and computational models. *Scientific Reports, 13*, 16120. https://doi.org/10.1038/s41598-023-43169-9

Bülthoff, I., & Newell, F. N. (2004). Categorical perception of sex occurs in familiar but not unfamiliar faces. *Visual Cognition, 11(7)*, 823–855. https://doi.org/10.1080/13506280444000012

Bülthoff, I., & Zhao, M. (2020). Personally familiar faces: Higher precision of memory for idiosyncratic than for categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46(7)*, 1309-1327. https://doi.org/10.1037/xlm0000784

Bülthoff, I., & Zhao, M. (2021). Average faces: How does the averaging process change faces physically and perceptually? *Cognition, 216*, 104867. https://doi.org/10.1016/j.cognition.2021.104867

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology, 51(3)*, 256–284.

Busey, T. A. (1998). Physical and psychological representations of faces: Evidence from morphing. Psychological Science, 9(6), 476–483. https://doi.org/10.1111/1467-9280.00088

Christenfeld, N. J. S., & Hill, E. A. . (1995). Whose baby are you? *Nature, 378*, 669. https://doi.org/10.1038/378669a0

Campanella, S., Hanoteau, C., Seron, X., Joassin, F., & Bruyer, R. (2003). Categorical perception of unfamiliar facial identities, the face-space metaphor, and the morphing technique. *Visual Cognition, 10(2)*, 129-156. https://doi.org/10.1080/713756676

Chauhan, V., Kotlewska, I., Tang, S., & Gobbini, M. I. (2020). How familiarity warps representation in the face space. *Journal of Vision, 20(7):18*. https://doi.org/10.1167/jov.20.7.18.

Costen, N.P., Parker, D.M. & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics 58,* 602–612. https://doi.org/10.3758/BF03213093

DeBruine, L. M., Smith, F. G., Jones, B. C., Roberts, S. C., Petrie, M., & Spector, T. D. (2009). Kin recognition signals in adult faces. *Vision Research, 49*, 38–43. https://doi.org/10.1016/j.visres.2008.09.025

Dobs, K., Bülthoff, I., & Schultz, J. (2016). Identity information content depends on the type of facial movement. Scientific Reports, 6, 34301. http://dx.doi.org/10.1038/srep34301

Dunn, J. C. (2010). How to fit models of recognition memory data using maximum likelihood. *International Journal of Psychological Research, 3(1)*, 140-149. https://doi.org/10.21500/20112084.859

Gilad, S., Meng, M., & Sinha, P. (2009). Role of ordinal contrast relationships in face encoding. *Proceedings of the National Academy of Sciences of the United States of America, 106(13),* 5353-5358. https://doi.org/10.1073/pnas.0812396106

Gilad-Gutnick, S., Yovel, G., & Sinha, P. (2012). Recognizing degraded faces: The contribution of configural and featural cues. *Perception, 41,* 1497-1511. https://doi.org/10.1068/p7064

Hancock, P. J.B. (2021). Familiar faces as islands of expertise. *Cognition, 214*, 104765,

https://doi.org/10.1016/j.cognition.2021.104765.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4(9)*, 330–337. https://doi.org/10.1016/S1364-6613(00)01519-9

Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science, 17(6)*, 485-492.

Jiang, F., Blanz, V., & O'Toole, A. J. (2009). Three-Dimensional Information in Face Representations Revealed by Identity Aftereffects. *Psychological Science, 20(3)*, 318-325. https://doi.org/10.1111/j.1467-9280.2009.02285.x

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17(5)*, 577–596. https://doi.org/10.1080/09658210902976969

Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers’ faces: Effects of the degree of relatedness. *Proceedings of the Royal Society B: Biological Sciences, 276(1670)*, 3193–3200. https://doi.org/10.1098/rspb.2009.0677

Kikutani, M., Roberson, D. & Hanley, J.R. (2008). What’s in the name? Categorical perception for unfamiliar faces can occur through labeling. *Psychonomic Bulletin & Review 15*, 787–794. https://doi.org/10.3758/PBR.15.4.787

Kikutani, M., Roberson, D., & Hanley, J. R. (2010). Categorical perception for unfamiliar faces: The effect of covert and overt face learning. *Psychological Science, 21(6)*, 865-872. https://doi.org/10.1177/0956797610371964

Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., Vandermalsburg, C., Wurtz, R. P., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers, 42*, 300–311.

Lander, K., Bruce, V. and Hill, H. (2001), Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology, 15*, 101-116. https://doi.org/10.1002/1099-0720(200101/02)15:1<101::AID-ACP697>3.0.CO;2-7

Leopold, D. A., O’Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4(1),* 89–94. https://doi.org/10.1038/82947

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6)*, 1364–1382. https://doi.org/10.1037/0278-7393.22.6.1364

Levin, D. T., & Beale, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception & Psychophysics, 62(2)*, 386–401. https://doi.org/10.3758/BF03205558

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience, 8(10)*, 1386–1390. https://doi.org/10.1038/nn1538

Maloney, L. T., & Dal Martello, M. F. (2006). Kin recognition and the perceived facial similarity of children. *Journal of Vision, 6*, 1047–1056.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27(3)*, 573–599. https://doi.org/10.1037/0096-1523.27.3.573

McLain, D. K., Setters, D., Moulton, M. P., & Pratt, A. E. (2000). Ascription of resemblance of newborns by parents and nonrelatives. *Evolution and Human Behavior, 21*, 11–23.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34(4)*, 865–876. https://doi.org/10.3758/BF03193433

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology, 102*, 726-747. https://doi.org/10.1111/j.2044-8295.2011.02053.x

Oda, R., Matsumoto-Oda, A., & Kurashima, O. (2005). Effects of belief in genetic relatedness on resemblance judgments by Japanese raters. *Evolution and Human Behavior, 26(5)*, 441–450. https://doi.org/10.1016/j.evolhumbehav.2005.04.002

Ramon, M., & Gobbini, M. I. (2018). Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition, 26(3)*, 179–195. https://doi.org/10.1080/13506285.2017.1405134

Ramon, M., Vizioli, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *Proceedings of the National Academy of Sciences of the United States of America, 112(35)*, E4835-E4844. https://doi.org/10.1073/pnas.1414929112

Robertson, D. J., Kramer, R. S. S., Burton, A. M. (2017). Fraudulent ID using face morphs: Experiments on human and automatic recognition. *PLOS ONE 12(3)*: e0173319. https://doi.org/10.1371/journal.pone.0173319

Rossion, B. (2002). Is sex categorization from faces really parallel to face recognition? *Visual Cognition, 9(8)*, 1003–1020. https://doi.org/10.1080/13506280143000485

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience, 8(1)*, 107-113.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research, 47(2)*, 157-165. https://doi.org/10.1016/j.visres.2006.11.002

Schwartz, L., & Yovel, G. (2019). Independent contribution of perceptual experience and social cognition to face recognition. *Cognition, 183*, 131-138. https://doi.org/10.1016/j.cognition.2018.11.003

Tanaka, J. Giles, M., Kremen, S., Simon, V. (1998). Mapping attractor fields in face space: the atypicality bias in face recognition. *Cognition, 68(3)*, 199-220. https://doi.org/10.1016/S0010-0277(98)00048-1.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25(4)*, 1016-1035.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36(12)*, 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 43(2)*, 161–204. https://doi.org/10.1080/14640749108400966

Watier, N. N., & Collin, C. A. (2009). Effects of familiarity on spatial frequency thresholds for face matching. *Perception, 38(10)*, 1497–1507. https://doi.org/10.1068/p6396

Wilson, H. R., Loffler, G., & Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research, 42*, 2909–2923.

Young, A. W., & Burton, A. M. (2018). Are We Face Experts? *Trends in Cognitive Sciences, 22(2)*, 100-110. https://doi.org/10.1016/j.tics.2017.11.007

Yue, X., Biederman, I., Mangini, M. C., Malsburg, C. Von Der, & Amir, O. (2012). Predicting the psychophysical similarity of faces and non-face complex shapes by image-based measures. *Vision Research, 55*, 41–46. https://doi.org/10.1016/j.visres.2011.12.012