THESIS

FAMILIARITY-DETECTION FROM DIFFERENT FACIAL FEATURE-TYPES: IS THE WHOLE GREATER THAN THE SUM OF ITS PARTS?

Submitted by

Brooke N. Carlaw

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2023

Master's Committee:

Advisor: Anne Cleary

Matthew Rhodes
Michael Thomas
Nathaniel Blanchard

ABSTRACT

FAMILIARITY-DETECTION FROM DIFFERENT FACIAL FEATURE-TYPES: IS THE WHOLE GREATER THAN THE SUM OF ITS PARTS?

Prior research indicates that perceived familiarity with a cue during cued recall failure can be systematically increased based on the amount of feature overlap between that cue and studied items in memory (Huebert et al., 2022; McNeely-White et al., 2021, Ryals & Cleary, 2012). However, these studies used word or musical stimuli. Faces represent a special class of stimuli, as evidence suggests that unlike other types of stimuli (such as word or musical stimuli), faces may be primarily processed in a holistic fashion. A recent study demonstrated that even when a person's identity was prevented by the presence of a facial occlusion like a surgical mask or sunglasses, familiarity-detection with the occluded face could still occur, suggesting that holistic processing was not a requirement for facial familiarity-detection (Carlaw et al., 2022). However, some researchers have suggested that although faces *can* be decomposed into component parts when partially occluded, when faces are presented unoccluded in their entirety, the holistic face processing system may then be obligatory (Manley et al., 2019). The present study suggests that this is not the case. Isolating specific feature types at *encoding* through partial occlusion of faces at study (via a surgical mask or sunglasses), then embedding those familiarized feature sets in otherwise novel whole faces at test, systematically and combinedly increased the perceived familiarity of the otherwise novel whole faces. These results suggest that even whole faces are processed as sets of component parts.

TABLE OF CONTENTS

Chapter 1 – Introduction

Familiarity

Familiarity is the sense of recognition or feeling that one has previously encountered something before. It can manifest as a sense of recognition when encountering a situation, despite failing to recall anything specific about a past encounter with that situation. The most common example of the experience of familiarity is that which occurs is with faces, as illustrated in the following quote by Rajaram (1993):

> *"There are times when we meet someone on the street whom we met at a party a few days ago. Although we know that we met this person at the party, we may not remember actually meeting the person, or his/her name" (Rajaram, 1993, p. 90).*

It is also illustrated in the way that Yonelinas (2002) described the experience:

> *"…The common experience of recognizing a person as familiar but not being able to recollect who the person is or where they were previously encountered" (Yonelinas, 2002, p. 441).*

The common use of familiarity with faces to illustrate the subjective experience of familiarity is also exemplified by the "butcher-on-the-bus" phenomenon described in the memory literature, whereby the now famous example of recognizing one's butcher outside of the usual context of the butcher shop (i.e., on the town bus) is used to illustrate the sensation of familiarity (see Brown, 2020; Mandler, 1991; MacLeod, 2020; Yovel & Paller, 2004).

Familiarity can be distinguished from the process of recollection (see Yonelinas, 2002, for a review). Whereas recognizing solely based on familiarity involves a mere sense or feeling of recognition, recognizing based on recollection involves calling to mind specifics about the prior experience in question (such as the identity of the person or the context in which that person was previously encountered).

Although faces are the most common example of familiarity-detection in day-to-day life, they are not the most investigated stimulus-type in laboratory studies of the process of familiarity; most commonly, researchers attempting to study familiarity have used words as stimuli (see Yonelinas, 2002, for a review). Thus, a glaring gap in the literature on familiarity concerns how familiarity occurs with faces. The present study aims to address this gap by investigating the mechanisms behind familiarity-detection with faces.

Mechanisms of Familiarity-Detection

One method of studying familiarity has involved investigating recognition without identification (RWI). RWI is the finding that, when only partial information is available in the test cue and identification is impaired, there is often still a sense of recognition (i.e., familiarity) (Peynircioğlu, 1990). Peynircioğlu found that when word fragments (e.g., *R _ _ ND _ _ P*) could not be identified, there were higher recognition ratings given to fragments of studied words (e.g., *RAINDROP*) than to fragments of unstudied words (e.g., *VERMOUTH*). This finding has since been extended to other forms of stimuli and other types of isolated features, including fragmented phonemes in spoken word fragments (Cleary et al., 2007), geometric shapes in picture fragments (Cleary et al., 2004), rapidly flashed masked pictures (Langley et al., 2008), and isolated pitch and rhythm information from piano pieces (Kostic & Cleary, 2009; McNeely-White et al., 2021).

The fact that RWI is most commonly elicited using isolated features of recently and non-recently presented stimuli makes it conducive to exploring familiarity within the context of global matching models of familiarity-detection. Global matching models specify that the familiarity signal, which can vary in intensity, arises from the degree of match between the features present in a test item and the features stored in memory from a study list (see Clark & Gronlund, 1996, for a review). An example is the MINERVA 2 Model (Hintzman, 1988).

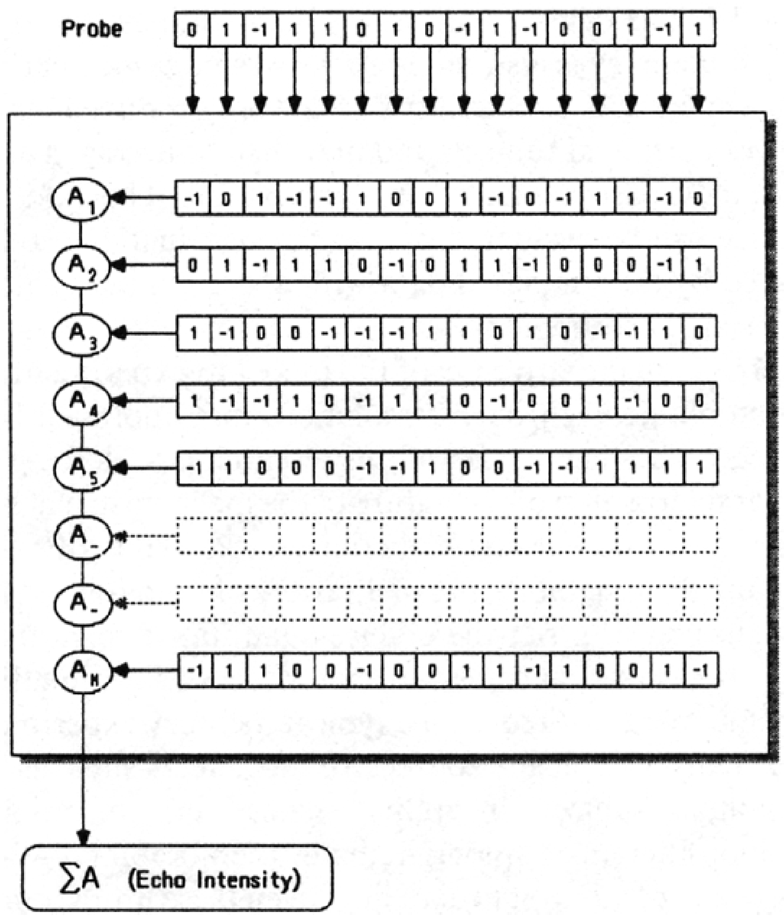

Figure 1. Hintzman's (1988) MINERVA 2 Model.

As shown in Figure 1, the MINERVA 2 Model assumes that each study list item is laid down in memory as a vector of features. Mathematically, an assessment of the degree of match between any given memory trace's features and the features present in the test item is computed (the A values in Figure 1), then each such assessment for every memory trace is summed to produce the familiarity signal (shown as Echo Intensity in Figure 1). The Echo Intensity value will be higher with a higher degree of feature match among a large number of memory traces (hence the term *global* matching model).

The RWI phenomenon is easy to explain within this type of model, as it can be accounted for by assuming that the limited features present in the test item enable enough of a match to features stored in recent memory to enable discrimination between unidentified studied and unidentified unstudied items. In fact, recently, researchers have attempted to systematically increase or decrease the level of perceived familiarity based on the degree of feature overlap between a test cue and items in memory during instances of cued recall failure (Cleary et al., 2016, Huebert et al., 2022; McNeely-White et al., 2021, Ryals & Cleary, 2012).

Ryals and Cleary (2012) used a variation of the RWI method known as recognition without cued recall (RWCR; Cleary, 2004) to evaluate whether systematically increasing the feature overlap in cues for studied words would correspondingly increase the level of perceived familiarity with the test cue in the absence of recall. When participants were unable to recall a studied word (e.g., *PITCHFORK)* in response to a cue (e.g., *POTCHBORK*) they reported higher ratings of familiarity for test cues that graphemically resembled a studied word than test cues with no resemblance to a studied word. Ryals and Cleary additionally found that increasing the feature overlap by studying four similar words to the test cue (e.g., *PITCHFORK, PULLCORK, PATCHWORK, POCKETBOOK*) systematically increased the perceived familiarity for the resembling cue (*POTCHBORK*) at test, even in the absence of recalling any of the studied similar words (e.g., *PATCHWORK*). These results suggest that the RWCR phenomenon can also be explained within the featuring matching framework of Hintzman's (1988) MINERVA 2 Model. Huebert et al. (2022) replicated this pattern while examining familiarity's role in prompting illusory recollection, and Cleary et al. (2016) found a similar pattern with different feature-type (semantic feature overlap).

In another study, McNeely-White and colleagues (2021) evaluated the feature-matching familiarity process of isolated musical features (rhythm and pitch) by having participants study those isolated feature types and later testing participants on the whole song clip. Specifically, in an encoding phase, participants studied isolated components (studied rhythm-only, pitch-only, or rhythm and pitch across separate isolated instances) of well-known piano pieces (e.g., children's melodies). For each song clip presented at study, the participant attempted to identify the clip (by typing the song's name). Then, in a test phase, each test trial contained presentation of the whole piano song clip from which those isolated study phase features had potentially come. Participants indicated if the musical piece was familiar ("yes, it is familiar" or "no, it is not familiar"), and answered an identification prompt (typed in the name of the song if they knew it). McNeely-White and colleagues found that participants systematically exhibited higher probabilities of reported familiarity for an entire piano test clip (e.g., Mary Had a Little Lamb) if they studied features of the song in isolation (e.g., rhythm or pitch) than if they had not. Additionally, if the test song clip had embedded within in it two different feature types that had been studied separately in isolation in the encoding phase (rhythm + pitch), the test song clip was found to be more familiar than if it had embedded within it only one (e.g., rhythm *or* pitch). This familiarity increase was additive (equivalent to adding together the level of familiarity increase from studying rhythm alone and the familiarity increase from studying pitch alone), which was shown by the fact that there was no difference in familiarity for the rhythm + pitch condition compared to the expected value obtained by adding together the (rhythm-only) + (pitch-only) conditions. This is consistent with how feature match indicators from different memory traces should generally combine to produce the intensity of a global familiarity signal in the MINERVA 2 model (Figure 1). Thus, encoding different types of musical features (rhythm and pitch) across different instances combinedly work to boost the familiarity of a later whole in which they are embedded, as would be predicted by the MINERVA 2 model of familiarity signal computation.

The present study is concerned with whether facial features studied in isolation similarly operate combinedly to boost the familiarity of a whole in which they are embedded. As previously noted, mechanisms of familiarity-detection with faces are less well-understood. Although faces are the most common example of familiarity in the literature, as evidenced by articles referring to the butcher-on-the-bus phenomenon (Brown, 2020; MacLeod, 2020; Mandler, 1991), they are not typically the stimulus of choice in studies of familiarity. Importantly, facial feature-types may behave differently than the aforementioned feature-types in familiarity signal computation. There is reason to suspect that faces may not be decomposed into constituent features in the same the way as other types of stimuli.

The Uniqueness of Faces

Faces may present a special case for feature matching theories of familiarity-detection and may even represent an exception to the general rule that items and events are segmented into separable features in memory. For instance, unlike other types of stimuli (such as words or music), faces may be perceived as a whole rather than broken-down into individual features (Tanaka & Simonyi, 2016). The intricacy and uniqueness of processing faces was described by Galton (1883):

> *"The differences in human features must be reckoned great, inasmuch as they enable us to distinguish a single known face among those of thousands of strangers, though they are mostly too minute for measurement. At the same time, they are exceedingly numerous. The general expression of a face is the sum of a multitude of small details, which are viewed in such rapid succession that we seem to perceive them all at a single glance" (Galton, 1883, p. 3).*

It is well-known that there are two components that make up the processing of a face: Featural and configural representations. Featural representations are the “parts” of the face (i.e., eyes, glabella, nose, mouth etc.) whereas the configural representation is the distance between those features (Mondloch et al., 2002; Tanaka & Simonyi, 2016). It is theorized that faces are configurally represented as a whole, whereby the entire makeup of the face is processed. This contrasts with the understanding of other forms of stimuli (e.g., words, music), which are thought to be processed by being decomposed into constituent features (e.g., letters, letter position, or rhythm, pitch, tone) rather than processed holistically (McNeely-White et al., 2021; Huebert & Cleary, 2022).

Holistic face processing can be characterized as being a situation in which the whole face is processed over its parts (Burton et al., 2015). Holistic face processing is typically tested using tasks such as the face inversion task (Yin, 1969), the face composite task (Young et al., 1987), and the part/whole task (Tanaka & Farah, 1993; see Bruyer, 2011, for a review). Using the face inversion task, Yin (1969) found that if a face is presented upside-down this reduces the ability to process and thereby recognize that face. However, when inverted object stimuli were presented, participants were better able to recognize them compared to inverted faces, consistent with the view that the way faces are processed is a special case that differs from other forms of stimuli like objects (see Valentine, 1988, for a review). Yin’s (1969) findings also suggested that configural representation of the features are primary to face identification ability, meaning that the distance between the features and how they are positioned on the face is a component of face processing.

Young and colleagues (1987) created a face composite task by taking the top half of one famous face and the bottom of another famous face and combining them to make one face. In doing so, they made a novel face from the two famous faces. Participants were less accurate and slower at correctly identifying the cued top or bottom halves in this case than if the halves were misaligned or if the combined face was inverted. The authors attributed this effect to the importance of using configural information to discriminate among faces (Young et al., 2018).

Tanaka and Farah (1993) created the part/whole face task by having the participant learn name-face associations and used a force-choice response procedure in which they presented the face part in isolation (e.g., a studied nose vs new nose as the options), as a whole (e.g., with the studied nose or a new nose), or as part of a whole with a new configuration or old configuration (e.g., changing the distance of the eyes). Tanaka and Farah found face superiority effects such that if the faces were studied as a whole, then recognition ability was more accurate if the test faces were also presented as a whole with the old configuration, rather than in isolation or with the new configuration. This is evidence in support of holistic processing of faces, demonstrating the importance of the configural representation, not just the features, for recognition ability.

Though holistic processing of faces is widely accepted, cognitive psychologists discount the notion that face processing relies solely on configuration, as features are also important (Cabeza & Kato, 2000; Tanaka & Simonyi, 2016). In fact, Tanaka and Simonyi concluded that on the holistic-analytic continuum (Farah, 1992), faces lie on the end of the range where there is a greater proportion of holistic processing than the part representations—*but* faces also contain a small proportion of part representations. In other words, faces are mainly processed holistically but *can* be broken-down into parts. However, it is less understood under what circumstances faces may be more likely to be broken down into parts. Some recent evidence suggests that a partial occlusion (like a mask) can cause a face to be primarily processed in terms of its parts rather than the whole (Manley et al., 2019). Still, the question of whether faces are (or can be) broken down into constituent parts for familiarity-detection remained relatively unexplored until recently (Carlaw et al., 2022). With the recent advent of the COVID-19 pandemic, a slew of new research aimed at understanding the impact of surgical face masks on face processing has shed some light on how occlusion of certain facial feature types (namely the nose and mouth information) affects face processing. The present study uses this basic method of facial occlusion at *encoding* to investigate the hypothesis that parts-processing may be involved in recognizing whole faces.

Does Configural Information Matter?

Though prior research suggests that configuration is an important factor for recognition (Tanaka & Farah, 1993; Yin, 1969; Young et al., 1987), it is unclear what distances of features matter. In addition, when the configuration of faces is changed, recognition ability is still intact for familiar faces (Burton et al., 2015). This is evidenced from prior research, such as Sandford and Burton's (2014) study where they changed the configuration of the faces by stretching the images. Participants were provided the stretched images of familiar (celebrity) and unfamiliar faces and asked to resize the image until it looked right. They found that participants struggled to correctly resize the images for familiar faces and, in one of the experiments, did so even more poorly compared to unfamiliar faces. These findings suggest that when configuration is altered for familiar faces, recognition ability is not hindered. This seems to point toward the notion that not as much weight should be on the configural representation for face recognition ability as has previously been suggested (Burton et al., 2015).

Differences in Facial Features

It is important to note that some features may hold more value than others in processing of faces. By using a principal component analysis for faces, Diego-Mas and colleagues (2020) found that for *perceived* facial traits of neutral faces, the greatest effects were for the eyes overall. However, other internal features also held value in perceiving different traits; for example, the mouth influenced happy, angry, or feminine traits. Furthermore, face structure seemed to also have an effect; for example, the jaw had a large effect on happy and angry traits (Diego-Mas et al., 2020). Eyes are also critical to the threat superiority effect (Fox & Damjanovic, 2006), the finding that threatening facial expressions are processed more quickly than neutral or nonthreatening expressions—also known as the "face in the crowd effect" (Fox & Damjanovic, 2006; Hansen & Hansen, 1988; Pinkham et al., 2010). Thus, prior research suggests that the eyes play a significant role in face processing and are used as a cue for perceiving traits of a face (Diego-Mas et al., 2020; Fox & Damjanovic, 2006).

Effects of Obstructions on the Face

Manley and colleagues (2019) investigated the impact of a perpetrator wearing a ski mask eyewitnesses' ability to correctly identify them later in a line-up. In a series of experiments, they had participants either encode full faces or partial faces and retrieve full faces or partial faces. For the partial-face conditions, the faces were either edited so the eyes were only presented, or the faces wore ski masks with just the eye information available. Manley and colleagues theorized that identification would be based upon transfer appropriate processing, suggesting that identification accuracy would be higher if the conditions in encoding matched the conditions at test compared to if the conditions did not match. They found that this theory was supported. That is, if participants only encoded the eye information (i.e., an image of a partial face or an image of a face wearing a ski mask), then they were more likely to correctly identify the studied face in a lineup if only the eye information was available (i.e., partial face or wearing a ski mask). Participants were better able to discriminate the target face from foils if it was only the eyes available as opposed to a lineup presented with full faces. Moreover, if participants studied full faces, they had better discriminability for the target in full face lineups compared to partial face/ski mask lineups. For target-absent lineups, if participants encoded a masked face, they were more likely to correctly reject the lineup of masked faces compared to full faces. But Manley et al. also found that when participants encoded a full face, they were also more likely to correctly reject the masked face lineup compared to a full-face lineup. The authors noted that this inconsistency may be attributed to participants being more conservative when much of the face information is not available (the forehead, nose, mouth, chin, etc.). These findings support the notion that faces tend to be processed holistically by default but *can be* processed in terms of features, including when doing so is necessary due to a facial occlusion like a mask.

Due to the COVID-19 pandemic, in recent years researchers have investigated how wearing a surgical face mask impacts the face processing (Carragher & Hancock, 2020; Freud et al., 2020). Carragher and Hancock (2020) created a series of face matching tasks with or without surgical face masks and found that the presence of a surgical face mask significantly impaired the ability to match those faces, as participants were less likely to match a face wearing a surgical face mask to the same face without a mask. They found this impairment for both familiar faces (celebrities) and unfamiliar faces (non-famous models), suggesting that regardless of whether a face is previously known or not, the presence of a face mask disrupts processing. Similarly with novel faces, Freud and colleagues (2020) evaluated whether face masks disrupt the ability to process that face using an adaption of the Cambridge Face Memory Test, a measure for face recognition ability. Freud and colleagues (2022) also found in a 20-month cross-sectional study and 12-month longitudinal study using the adapted Cambridge Face Memory Test that recognition impairment did not improve over the course of the COVID-19 pandemic, when mask wearing was common. This suggests that there is an inability to adapt to visual changes in faces caused by the regular occlusion of wearing a surgical mask. Given these studies, there is compelling evidence that the presence of a surgical face mask on both previously known and novel faces disrupts face processing. That said, it still may be the case that people *can* rely on parts-processing of facial information in at least some circumstances (e.g., Manley et al., 2019), and recent research suggests that one of those circumstances might be for familiarity-detection during instances of facial occlusion (Carlaw et al., 2022).

Facial Familiarity-detection

Although faces are not typically chosen as the stimuli for studying familiarity, some studies do suggest that familiarity plays a role in face recognition. For example, Rhodes and colleagues (2008) examined what factors contribute to older adults falsely recognizing pairs of faces that were previously studied as members of other pairs by using an associative recognition test. Most applicable to the current study, they examined whether familiarity is a contributor to these types of false alarm matches. Participants studied half of the face pairs once and the remaining face pairs four times (half of the face pairs were older adults and half were younger adults), then were tested on intact face pairs, rearranged face pairs (face pairs containing studied faces that had been previously studied separately as members of other pairs rather than together as a pair), and new face pairs (unstudied faces). The main finding was that both younger and older participants tended to falsely recognize the rearranged face pairs of those they previously studied in different pairs. They also found that by studying the face pair multiple times (four times vs. once) the rate of false alarms increased for the older adults. Thus, these findings suggest that for multiple presentations of faces, familiarity (which increases with the number of studied presentations) likely plays a major role in associative face recognition (over recollection of contextual details) given the increase in false alarms (Rhodes et al., 2008).

Recognition without Identification (RWI) of Faces

While most RWI studies involve list-learning paradigms, Cleary et al. (2013) took a unique approach aimed at more closely approximating the real-life experience of recognizing a face as familiar as opposed to novel. Specifically, they investigated RWI of faces by having participants view images of celebrity and non-celebrity faces whose identification had been prevented by a Gaussian monochromatic noise filter. To ensure that participants knew the celebrities, they were later presented with only the celebrity faces unfiltered and given the opportunity to identify them. Overall, when known celebrity faces (i.e., that were later identified without the filter) were unidentified through the filter, these celebrity faces received higher familiarity ratings compared to when the filtered faces were non-celebrity faces.

Carlaw et al. (2022) used the same celebrity face stimuli as Cleary and colleagues without the blur filter and instead added a surgical face mask or sunglasses to obstruct the respective facial feature types (mouth and nose or eyes), hence separating the face by features. Participants viewed masked (or sunglasses-covered) and hooded famous and non-famous faces and were asked to rate how familiar the person felt (*0 = Not at all familiar-seeming, 10 = Extremely familiar-seeming*). They were also asked if they could identify the person; if they could, they were instructed to name them or provide any identifying information they could recollect. Following this, participants viewed only the celebrity faces unmasked and unhooded and were again asked to identify the person. Carlaw and colleagues found that, in the absence of identification for masked or sunglasses-covered celebrity faces that were later identified unmasked (i.e., the participant knew the celebrity but was unable to name them or provide any information about them with the surgical mask or sunglasses and hood occlusions), familiarity ratings were higher for unidentified famous faces than similarly occluded non-famous faces. In other words, in the absence of face identification, familiarity-detection occurred for the information that was available (eyes, eyebrows, glabella, forehead or nose, mouth, chin).

Despite the finding that the decrement for identification was greater for known celebrity faces wearing sunglasses than surgical masks (identification was significantly more impaired by covering eye feature information compared to covering nose and mouth feature information), Carlaw et al. (2022) not only found a significant RWI effect in both cases of occlusion (surgical masks vs. sunglasses), but also found no difference in the magnitude of the RWI effect among faces wearing masks or sunglasses. This suggests that familiarity-detection might have been similarly driven by the different facial feature-types investigated. If this is the case, then familiarity-detection of faces might be comparable to other types of stimuli that are broken down into features or parts for familiarity signal computation, such as words (e.g., Huebert & Cleary, 2022) or music (Kostic & Cleary, 2009; McNeely-White et al., 2021). This makes it plausible that facial features might combine globally to produce an overall familiarity signal with a face, and the method of obstructing a face with a surgical mask or sunglasses presents a potentially useful means of investigating this.

Do Parts of Faces Combine Across Traces to Familiarize the Whole?

As previously described, McNeely-White et al. (2021) found that the musical feature types of rhythm and pitch combined across separately encoded memory traces to boost the perceived familiarity of a whole song clip in a manner consistent with the proposed operation of the familiarity signal computation in MINERVA 2 (see Figure 1). The present study used the facial occlusion method (of surgical mask vs. sunglasses occlusion) to examine whether the same might be true of facial feature sets.

Some clues to the possibility that faces might be broken down into constituent features that are combined for familiarity signal computation with a face can be found in some connections in the literature on facial attractiveness. For instance, faces containing the average of all facial features tend to be perceived as more attractive than faces with more distinctive features (Valentine et al., 2004). In addition, people will tend to fill in missing information (such as when a face is occluded by a surgical mask) with more attractive features (very likely to be average-seeming features), thus finding unmasked faces more unattractive than anticipated before they were unmasked (Pazhoohi & Kingstone, 2022). Note that, according to the MINERVA 2 Model illustrated in Figure 1, a face that is the average of all features should be more familiar than one that is less so, because the average will contain a higher degree of global match to all the features stored in memory than a face that has more distinct features and does not map onto as many faces in memory. Implicit in this idea is that facial features are averaged in the mind in order to contribute to perceived attractiveness in the first place. Given the finding the familiarity and positivity or liking are strongly correlated (e.g., Monin, 2003), it is possible that the averageness of a face drives both perceived attractiveness and familiarity, though this idea has not been directly explored.

Given the links between averageness of a face and perceived attractiveness, and the known correlation between positivity, liking and familiarity, it is reasonable to hypothesize that facial features might combine in memory analogously to how other stimulus feature-types have been shown to combine in memory (e.g., McNeely-White et al., 2021). If so, this is important for understanding face processing more generally, including fully understanding how feature-based memory processes might contribute not only to familiarity with faces, but also other perceptions of faces such as attractiveness or trustworthiness. Moreover, given the likely role of initial familiarity-detection in prompting memory search in the first place (Carlaw et al., 2022), understanding how familiarity-detection takes place with faces has broader implications for how people make decisions about faces and determine a person's identity.

The Present Study

The present study aims to determine if facial feature types combine across memory traces in ways predicted by global matching approaches to familiarity signal computation and in ways shown with other types of stimuli. For example, McNeely-White et al. (2021) found that a piano clip (e.g., Mary Had a Little Lamb) at test was perceived to be more familiar if a feature-type from it had been studied in isolation in an earlier study phase (e.g., the tapped-out rhythm of Mary Had a Little Lamb in a single note on a wood block instrument was heard in an earlier study phase). Interestingly, McNeely-White et al. observed that the perceived familiarity of the full piano clip at test was even greater when two different feature-types from the piece had been studied in isolation across separate instances in the study phase (i.e., the isolated rhythm and the isolated pitch information had been separately encoded) than when only one feature-type had been heard in the study phase. In short, when two different familiarized feature-types (rhythm and pitch) were embedded in the test song clip, the song clip seemed more familiar than when only one familiarized feature-type (e.g., rhythm only) had been heard in the study phase. Note that this is what the model depicted in Figure 1 would predict.

If faces are processed as sets of features when it comes to familiarity signal computation, as the finding that average faces are more attractive might suggest (Valentine et al., 2004), then one might predict a similar pattern for facial feature types in a paradigm analogous to that which was carried out by McNeely-White et al. (2021) but with faces instead of music. That is the aim of the proposed study. Analogously to how McNeely-White and colleagues examined musical features, the present study examined if familiarizing certain facial feature-types through different methods of facial occlusion will show evidence of a similar type of global matching process for the perceived familiarity of the faces in which these familiarized features are embedded. Specifically, I not only examined whether familiarizing a particular facial feature set (e.g., nose and mouth information in the sunglasses-only condition or eye and eyebrow information in the mask-only condition) would increase the perceived familiarity of a later whole face in which those features are embedded (relative to if none of its features had been earlier familiarized), but also examined whether familiarizing different facial feature types (e.g., nose and mouth information in one instance *and* eye and eyebrow information in another instance) across different study episodes (in a mask + sunglasses condition) would increase the perceived familiarity of the whole face in which these feature-types are embedded relative to familiarizing only one facial feature-type. Specifically, if participants study a face in which part of it is occluded by sunglasses and, in a separate instance, study a surgically masked face, is the latter full face perceived as more familiar than if it had only been studied with sunglasses or only with a surgical mask?

Carlaw et al. (2022) prior finding that, in the absence of identification of known-occluded faces, participants detected familiarity from the facial features available (e.g., eyes, glabella and forehead, or nose, mouth, and chin), suggests that partially occluded faces can be broken-down into feature types for familiarity-detection. Thus, we can examine whether separately studying the isolated feature types of the face can systemically boost the perceived familiarity for a whole face, and if so, if the increase occurs in a way that is additive (analogous to what McNeely-White et al., 2021, found with musical features), or if it occurs in a way that is either sub-additive or multiplicative.

The central question of the current study is if the familiarization of particular isolated facial feature sets (e.g., nose-mouth features or eye-eyebrow features) would lead to an increase in the perceived familiarity of a later whole face in which these familiarized features are embedded relative to if no feature sets had been familiarized. And, if so, whether separate exposures to each different isolated feature set would confer additional perceived familiarity with the whole face in which those features are embedded than if only one feature type had been viewed at encoding. If whole unoccluded faces are obligatorily processed holistically (e.g., Manley et al., 2019), then there may be no increase in perceived familiarity with a whole face with an increasing number of familiarized feature-types within the otherwise novel whole face (i.e., that was never presented as a whole to the person previously). That is, the whole face may seem just as familiar when no component features were studied as when one subset of features were studied (i.e., either nose-mouth features or eye-eyebrow features) as when a combination of feature-types (i.e., both nose-mouth features and eye-eyebrow features across separate encoding instances). If so, the resulting pattern across conditions might be:

(mask + sunglasses) = [(mask-only) = (sunglasses-only)] = unstudied faces

Although this pattern seems unlikely given that familiarity-detection with occluded faces that were presumably stored in memory as wholes has already been shown to occur (Carlaw et al., 2022), it remains an open question whether faces presented as wholes will receive a boost in their perceived familiarity from when separated components have been familiarized.

If like with musical stimuli (McNeely-White et al., 2021), exposure to isolated facial feature sets boosts the later perceived familiarity with the whole in which they are embedded and does so combinatorically across separate exposures to different feature sets that are later embedded within the same whole face (see Figure 2 below), then the familiarity rating pattern should be such that:

(mask + sunglasses) > [(mask-only) = (sunglasses-only)] > unstudied faces

Of note, the presumption that familiarity will be roughly equal in the mask only and the sunglasses only condition is based on the pattern found by Carlaw et al. (2022), which suggested that the level of perceived familiarity increase of the available features with the mask and sunglasses occlusions were equal. However, given that prior research suggests that eyes carry more value in the processing of faces than other features (Diego-Mas et al., 2020; Fox & Damjanovic, 2006), it is possible that the eye-eyebrow information will carry more weight in the computation of the familiarity signal than nose-mouth information. That said, although Carlaw et al. (2022) found that identification to known faces was hindered more when the eyes were covered (i.e., wearing sunglasses) than when the nose and mouth features were covered (with a surgical mask), they found that there was no difference in perceived familiarity in either case. Thus, it is likely that familiarity-detection from nose-mouth features vs. for eye-eyebrow features is approximately equal in intensity.

As mentioned, McNeely-White et al. (2021) found that the musical feature sets of rhythm and pitch combined across memory traces to boost the familiarity of whole piano piece in which they were embedded in a way that was *additive*. Thus, one potential pattern in the present study may be that the level of familiarity boost found for the mask + sunglasses condition will be *additive* in nature. Specifically, the level of increase would be roughly equal if we added together the level of familiarity boost found for the mask-only condition and the level of familiarity boost found for the sunglasses-only condition:

(masks + sunglasses) = [(mask-only) + (sunglasses-only)] > unstudied faces

This would be analogous to McNeely-White et al.'s (2021) findings with musical features, where they found that rhythm and pitch additively combined to boost the familiarity with the whole piano song clip in which they were embedded, consistent with the mechanisms of the MINERVA 2 model depicted in Figure 1.

However, given the aforementioned suggestions in the literature that face configuration and holistic processing are essential to processing of whole faces, there is reason to hypothesize that, unlike other types of stimuli (such as musical stimuli) where the whole may be equal to the sum of its parts (McNeely-White et al., 2021), with faces the whole may be greater than the sum of its parts. A way that this might manifest is as:

(mask + sunglasses) > [(mask-only) + (sunglasses-only)] > unstudied faces

Specifically, if the masks + sunglasses condition results in increased perceived familiarity, this would suggest that the presence of a co-occurring combination of previously familiarized feature-types carries more in formation value than simply adding the familiarized features' familiarity increase together. Thus, if it were significantly greater than what would be expected by simply adding the level of familiarity increase afforded by the mask-only condition with the level of familiarity increase afforded by the sunglasses-only condition, this would imply that having more than one previously familiarized feature-type within an otherwise novel whole face that was never before seen configurally as a whole carries additional weight in the computation of the familiarity signal for the whole. That is, there may be an "added boost" from co-occurring familiar information, particularly given that configural information is thought to be important to processing faces (e.g., Yin, 1969; Young et al., 1987). However, this would still predict that facial features add to the overall familiarity as well. Thus, the parts would be contributing to the overall perceived familiarity, but the whole would be greater than the sum of its parts.

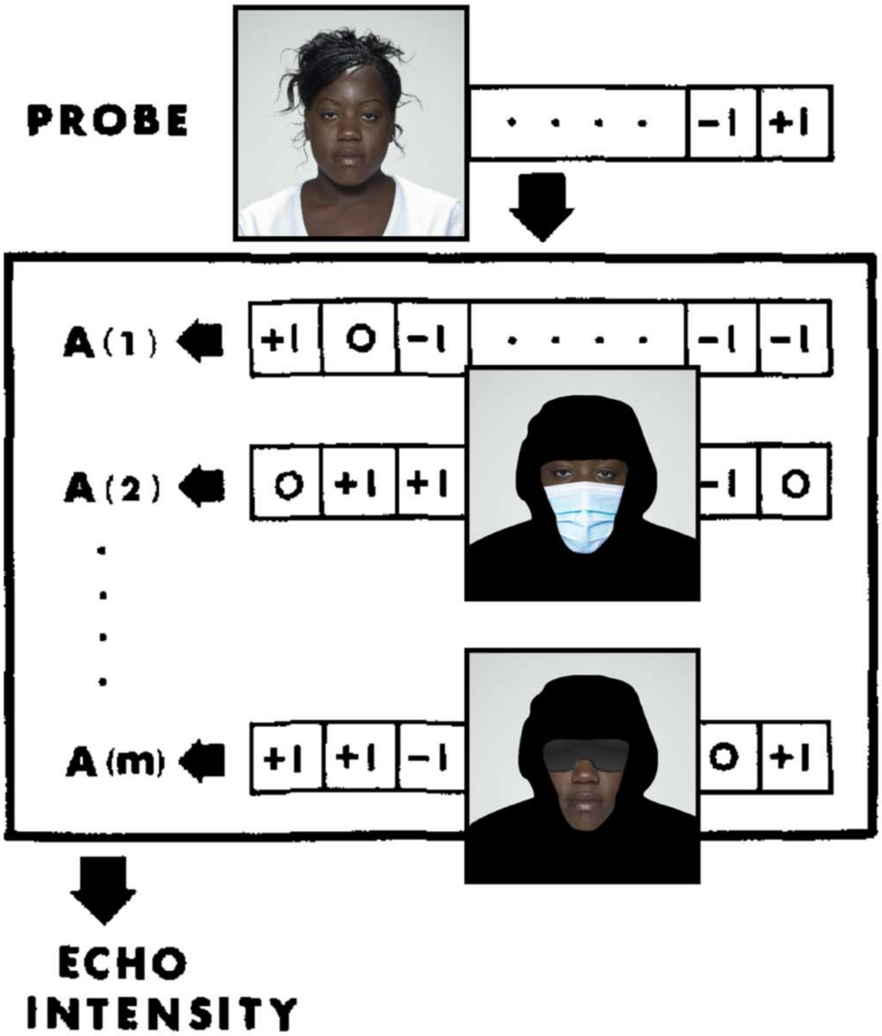

Figure 2. An example of how studying separate facial features would be imbedded into Hintzman's MINERVA 2 Model. Original face image from the Face Research Lab London Set: https://doi.org/10.6084/m9.figshare.5047666.v5

To more rigorously examine the question of whether whole unoccluded faces may launch an obligatory holistic processing system (e.g., Manley et al., 2019), we used completely novel faces rather than faces that were known to the participant for two reasons. First, Carlaw et al. (2022) showed that familiarity-detection with partially occluded unidentified but known faces could still occur (suggesting that facial features were extractable from memory representations resulting from prior whole face exposure). Thus, this leaves open the question of whether completely novel whole faces would tend to be processed holistically. Second, the use of novel faces from an existing face image pool enabled control over the facial positioning, lighting, and background of all the faces.

Chapter 2 – Experiment

Method

Participants

One hundred twenty-three participants were recruited from the Psychology Research pool at Colorado State University and received credit toward a course requirement for their in-person participation. A sensitivity analysis was run using G*Power, with the parameters set as the power being 95% (α =.05) for the sample size of 123; the analysis revealed that the effect size of .33 was sufficient to detect the magnitude of the effect. Therefore, the experiment was correctly powered.

For the current experiment, recruitment was limited to people who have normal or corrected-to-normal vision and hearing, and native-level fluency in English. Upon completion of the experiment, demographic information was collected such as age, gender identity, educational level, and ethnicity. The mean age of the participants was 19 years old (*SD* = 1.38), .44 participants identified as male, .52 identified as female, 1 participant identified as nonbinary or third gender, and 1 participant identified as a transgender man. The average educational level (*12 years = first year of college*) of participants was 12.5 years (*SD* = 1.26). The sample demographics consisted of .70 who identified as Caucasian/White (non-Hispanic), .03 Black (non-Hispanic), .02 Asian or Pacific Islander, .12 Hispanic, .09 who identified as multiple ethnicities and 1 participant who selected "Prefer not to say." Participants were also asked to indicate *Yes* or *No* if they are a native English speaker and if they selected *No*, they were prompted to indicate how many years they have spoken English. Four participants indicated they were not a native English speaker and spoke English for an average of 10 years. Due to research assistant error, demographic data was not collected for 1 participant.

Design

The experiment is a within-subjects design with three conditions: 1) single presentation of the novel faces (mask-only), 2) single presentation of sunglasses-covered novel faces (sunglasses-only), and 3) separate presentation of the opposite facial feature types of a novel face across the study phase, where the same face was presented once as a sunglasses-covered face and once as a masked face (masked + sunglasses).

Materials

The experiment was implemented in-person using E-Prime. The stimuli consisted of 80 images of novel faces that were selected from the Face Research Lab London Set (DeBruine & Jones, 2017) and saved as JPEG files. Each image was 1350 x 1350 pixels, full color with similar lighting and grey background. The faces in each image had the same pose and were in the same position in the image. In order to control for any potential differences in lighting, image contrast, or skin pigment, in addition to each subject analysis, they were also run with the items treated as subjects.

There were 80 total faces chosen for use in the experiment with 40 female and 40 male faces accounting for diversity of race and ethnicity in the stimulus set (12 faces of the images collected identified as Black, 9 identified as East Asian, 1 identified as East Asian/White, and 48 faces identified as White). A total of 240 variants of these face stimuli were used (80 of these were the unoccluded test faces; 80 were the surgical mask occluded versions of these faces for use at study, and 80 were the sunglasses-occluded versions of these faces also for use at study). Using Adobe Photoshop, the surgical mask, sunglasses, and hood images overlaid on the faces were selected from SearchPng (searchpng.com), Klipartz (Klipartz.com), and PNGIX (pngix.com) respectively and be downloaded as PNG files. The sunglasses and mask images were chosen based on there being little overlap of the skin between such features, in an attempt to control familiarization of parts of the face being exposed twice for the mask + sunglasses condition.

For the study phase stimuli, there were three facial feature-type conditions. To ensure that the familiarization would be strictly limited to the facial feature types of interest within each condition, each study phase face was made to be wearing a hood (so that hair information does not become familiarized in the encoding phase). For the study phase facial feature-type familiarization, each hooded face had two versions (see Figure 3 below). In one, the nose and mouth area was covered with a surgical mask (see the left-hand panel of Figure 3). In the other, the eye, glabella and eyebrow area were covered with sunglasses (see the middle panel of Figure 3). During the study phase of each of four blocks, a total of 15 different occluded faces were presented. Five of these were occluded with a surgical mask only; five were occluded with sunglasses only, and five appeared across two different study phase instances in each occlusion condition (i.e., the hooded face appeared in one instance of surgical mask only and one instance of sunglasses only across two different randomly ordered episodes). Thus, the study phase consisted of 20 total presentations.

For the test phase, the faces were presented completely unoccluded (i.e., no surgical mask, no sunglasses, and no hood), as in the right panel of Figure 3. In each block, 20 unoccluded test faces were shown (5 that had been studied with a mask only, 5 that had been studied with sunglasses only, 5 that had been studied once with a mask and once with sunglasses, and 5 that were not studied).

The 80 faces were divided into four study-test blocks. In each block, a 20-item randomly-ordered study list (5 faces occurring once masked and hooded; 5 occurring once sunglasses-covered and hooded, and 5 occurring twice—once masked and once sunglasses-covered) were followed by a 20-item test list. The test list contained the 15 studied faces unmasked along with 5 additional non-studied unmasked faces in a random order (see right-hand panel of Figure 3).

All analyses were performed using traditional null hypothesis significance testing and were run with IBM SPSS. In addition, Bayes Factors were calculated with JASP, and the prior classified was JZS (see Wagenmakers, 2007 for the classification recommendations). Cohen's *d* for repeated-measures was also used for the effect size calculations (drm; see Lakens, 2013).

Figure 3. An example of one face from the stimulus pool: The first image wearing a surgical mask and hood; the second wearing sunglasses and hood; third unmasked, and the sunglasses/hood are removed. Original image from the Face Research Lab London Set: https://doi.org/10.6084/m9.figshare.5047666.v5

Procedure

The experiment began with a welcome screen, instructing participants to press any key to continue (see Figure 4 below for a break-down of the experiment for each condition). Then there was a set of instructions (taken from Carlaw et al.'s (2022) second experiment with a few edits to the dialogue to fit the current experiment) on the screen for the study phase:

Welcome to this study on human memory!

In this study, one-at-a-time you will view faces of people who are wearing a medical mask and a hood or a pair of sunglasses and a hood.

Press any key to continue.

Once any key was pressed, the next set of instructions for the test portion appeared:

Once you have viewed all of the faces wearing masks, sunglasses and hoods, you will view a set of faces with no coverings or hood. For each face, you will be asked to rate how familiar the face seems using a scale of 0 to 10 where 0 means "not at all familiar-seeming" and 10 means "extremely familiar-seeming." Then, you will be asked if you feel that you can recall a prior instance of seeing each face at study by responding with "yes" or "no". You must give a familiarity rating between 0 and 10 and answer "yes" or "no" for each face.

Press any key to start the experiment.

Once a key was pressed, the randomly ordered partially occluded faces flashed by on the screen one at a time in the center of the screen for a duration of 2 s each. Once all 20 partially occluded faces were presented, a screen appeared with test instructions:

You will now complete the test phase of the experiment.
Please press ENTER to receive instructions.

Upon proceeding, the next set of instructions appeared:

You will now view a set of faces with no coverings or hood. For each face, you will be asked to rate how familiar the face seems using a scale of 0 to 10 where 0 means "not at all familiar-seeming" and 10 means "extremely familiar-seeming."

Then, you will be asked if you feel that you can recall a prior instance of seeing each face at study by responding with "yes" or "no". You must give a familiarity rating between 0 and 10 and answer "yes" or "no" for each face.

Press any key to continue.

Following the key press, each unhooded, unoccluded, fully visible face was randomly selected to appear one-at-a-time in the center top of the screen. Below each face, a series of prompts appeared. Participants were asked to indicate if the face feels familiar, utilizing Carlaw et al.'s (2022) prompt (*On a scale of 0-10, how familiar does this person feel? 0 = Not at all familiar-seeming, 10 = extremely familiar-seeming*) and typed the number below the prompt.

Participants were asked, *Do you feel that you can recall a prior instance of seeing this face in the study phase? Press "Y" for YES or press "N" for NO*. They again indicated their response and typed "Y" or "N". Once all test images were randomly presented and the prompts for each of the faces were completed, the second block was completed. Each block proceeded in the same manner as the first block, but with a new set of stimuli. In sum, each block had 20 faces at test (see Figure 4). At the end of the experiment, a screen appeared thanking the participants for completing the experiment. Following this, the participants completed a written demographics form. For the demographic questions, participants were asked to indicate their responses for their age, gender identity, educational level, ethnicity, and their proficiency in English language. After each of these questions were answered, the participant was debriefed and provided credit for their participation.

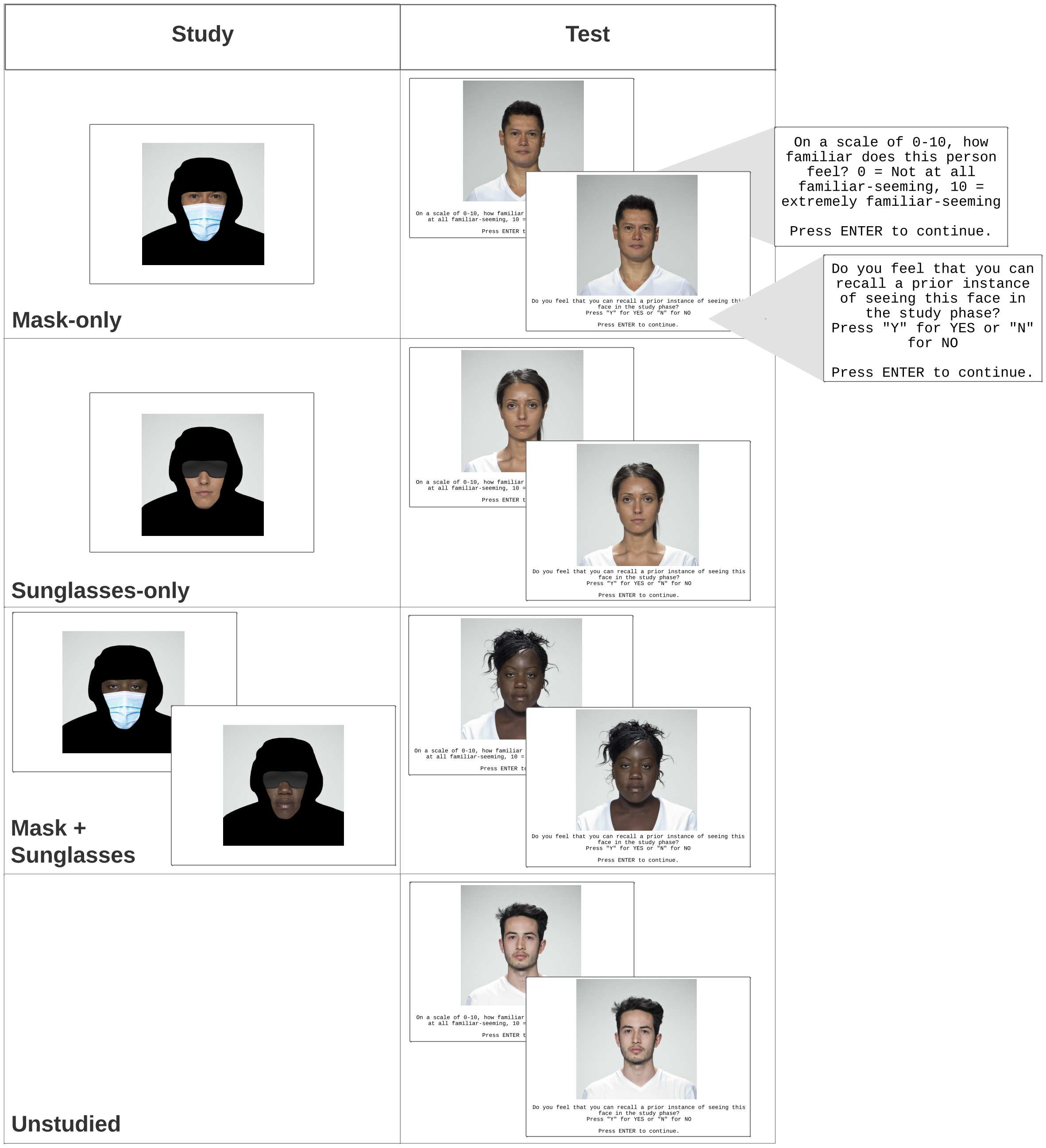

Figure 4. Visual illustration of the E-Prime experiment depicting each condition (mask-only, sunglasses-only, masked + sunglasses) and non-studied faces. Original face images from the Face Research Lab London Set: https://doi.org/10.6084/m9.figshare.5047666.v5

Chapter 3 – Results

Overall Familiarity Ratings

Analyses were performed for both subjects and items—in addition to the usual subject analyses, a version of each analysis was conducted in which items were treated as participants in the analysis. The items analyses were performed to hold constant the face itself (e.g., skin pigment, texture, lighting, contrast) across the experimental conditions. Also, given the unique intrinsic properties of each face, the configuration and features are exclusive to that face alone and thus having the items analyses will hold those variations constant. For the analysis of the overall face recognition data, a repeated-measures ANOVA performed on the average familiarity ratings (see Table 1, Figures 5 and 6 below) for each face condition (unstudied, mask-only, sunglasses-only, mask + sunglasses) revealed a significant main effect $F_{subjects}(3, 366) = 133.64$, $MSE = 97.52$, $p < .001$, $\eta_p^2 = .523$, $BF_{10} = 8.49 \times 10^{54}$; $F_{items}(3, 237) = 77.21$, $MSE = 62.45$, $p < .001$, $\eta_p^2 = .494$, $BF_{10} = 1.09 \times 10^{32}$.

Table 1. Average familiarity ratings for each type of feature studied for subjects and items.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	3.79	1.37	3.75	.96
Mask	5.15	1.21	5.13	1.18
Sunglasses	5.13	1.32	5.09	1.12
Mask + Sunglasses	5.93	1.22	5.88	1.20

Planned comparisons suggested that, overall, participants gave higher familiarity ratings to whole faces presented at test that had been masked at study than to unstudied faces, $t_{subjects}(122) = -11.74$, $SE = .12$, $p < .001$, $d = -1.05$, $BF_{10} = 4.98 \times 10^{18}$; $t_{items}(79) = -8.88$, $SE = .16$, $p < .001$, $d = -1.28$ $BF_{10} = 4.65 \times 10^{10}$. Participants also gave higher familiarity ratings to whole faces at test that were studied wearing sunglasses than to unstudied faces, $t_{subjects}(122) = -12.08$, $SE = .11$, $p < .001$, $d = -1.0$, $BF_{10} = 3.13 \times 10^{19}$; $t_{items}(79) = -11.02$, $SE = .12$, $p < .001$, $d = -1.28$, $BF_{10} = 4.90 \times 10^{14}$. When comparing the faces that were masked at study to faces wearing a pair of sunglasses at study, there was no significant difference, $t_{subjects}(122) = .12$, $SE = .10$, $p = .906$, $d = .02$ $BF_{01} = 9.91$; $t_{items}(79) = .21$, $SE = .17$, $p = .838$, $d = .04$, $BF_{01} = 7.96$. This lack of difference suggests that the nose-mouth features, when isolated, and the eye/eyebrow features, when isolated, boosted familiarity levels by approximately the same amount. In other words, they carried the same weight in terms of the level of familiarity increase shown later with the whole face in which they were embedded. Participants also gave higher familiarity ratings to whole test faces in the mask + sunglasses condition, which had both feature types (nose-mouth and eye/eyebrows) studied across two separate instances, once wearing a mask (eyes/eyebrows) and another wearing sunglasses (nose-mouth) than studying the face with a mask-only, $t_{subjects}(122) = -9.23$, $SE = .09$, $p < .001$, $d = -.64$, $BF_{10} = 5.92 \times 10^{12}$; $t_{items}(79) = 7.40$, $SE = .10$, $p < .001$, $d = -.63$, $BF_{10} = 7.61 \times 10^{7}$ or than studying the face with sunglasses-only, $t_{subjects}(122) = -7.85$, $SE = .10$, $p < .001$, $d = -.63$, $BF_{10} = 4.23 \times 10^{9}$; $t_{items}(79) = 5.45$, $SE = .14$, $p < .001$, $d = -.86$, $BF_{10} = 25337.38$.

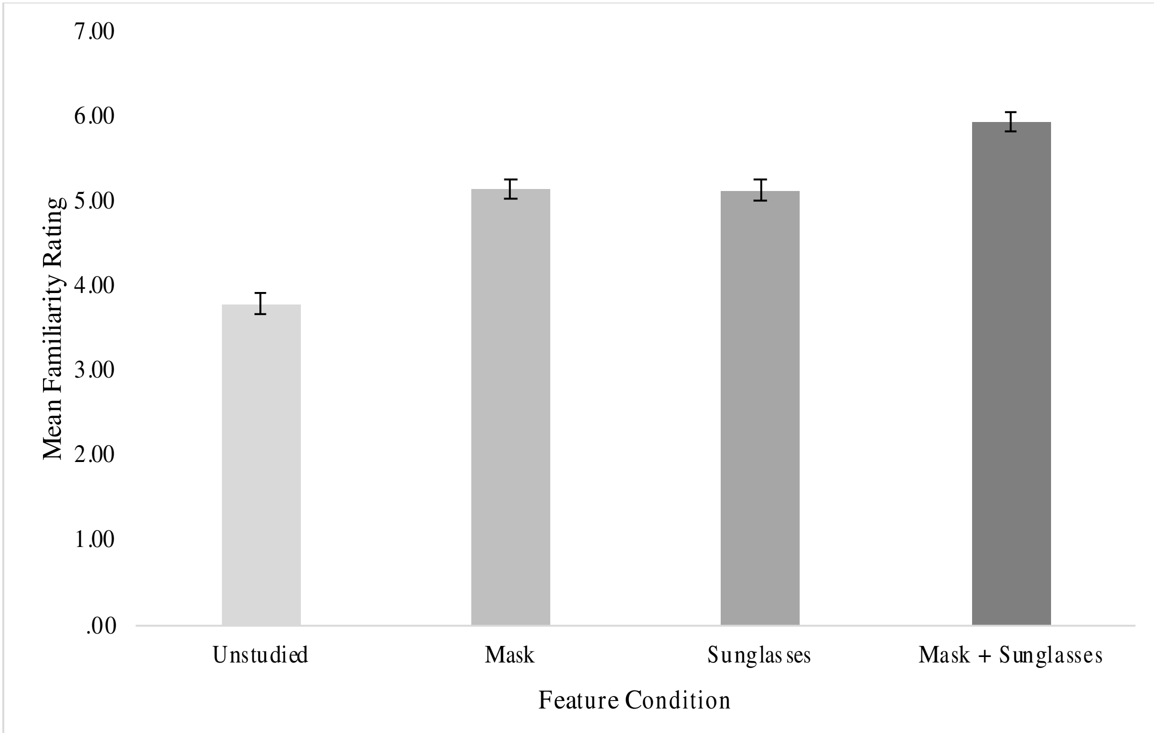

Figure 5. Average familiarity ratings for each type of feature studied for subjects.

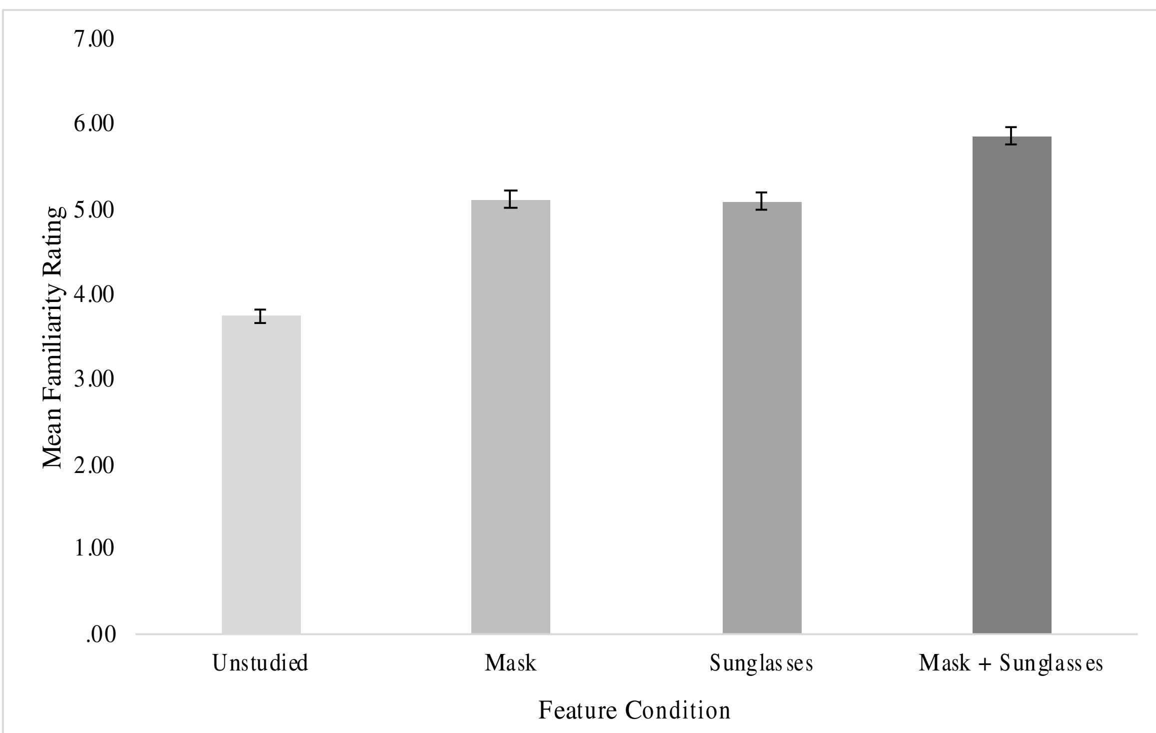

Figure 6. Average familiarity ratings for each type of feature studied across items.

To examine the level of increase in familiarity ratings (see Table 2 below) across the different feature-familiarization conditions, a one-sample *t*-test was run for the average level of increase for mask + sunglasses condition ($M = 2.14$) against 2.71 (the level of increase shown by adding together the increase from the mask-only and sunglasses-only conditions relative to the unstudied baseline condition). This analysis revealed a significant difference, $t_{subjects}(122) = -4.15$, $SE = .14$, $p < .001$, $BF_{10} = 264.81$. For the items, the one-sample *t*-test compared 2.13 against 2.72, which revealed a significant difference, $t_{items}(79) = -3.78$, $SE = .16$, $p < .001$, $BF_{10} = 73.43$. This overall pattern suggests a sub-additive increase in familiarity for the mask + sunglasses condition.

Table 2. Mean level of familiarity-rating increase for each type of feature studied for subjects and items.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	-	-	-	-
Mask	1.36	1.28	1.38	1.39
Sunglasses	1.35	1.24	1.34	1.09
Mask + Sunglasses	2.14	1.51	2.13	1.40

The Subjective Sense of Recall

Since all of the faces used in the present study were novel, participants did not have pre-experimental knowledge of these faces to know a name for each face. Therefore, unlike in Carlaw et al. (2022), we could not objectively assess participants' ability to consciously identify faces through the occlusions at study. Instead, participants were essentially learning novel faces through first (and sometimes second, in the mask + sunglasses) exposure in a situation of partial occlusion, without having had prior exposure to the whole face. As an alternative to being able to examine face identification ability at encoding, we examined participants' subjective sense of whether they thought they were recalling the exact face from study when viewing the whole face at test. We added this primarily to examine 1) how often participants subjectively felt that they were recollecting the specific occluded face corresponding to the otherwise novel whole test face and 2) how familiarity ratings varied as a function of participants' subjective metacognitive impressions of their memories.

Although this particular method is now considered an antiquated method of trying to separate out the contributions of recollection versus familiarity that has long been described by researchers as problematic (e.g., Rotello et al., 2005; Wais et al., 2008), recent research on the possibility of familiarity-driven illusory recollection/commission errors (Carlaw et al., 2022; Huebert et al., 2022) suggests that strong familiarity might subjectively be *interpreted* as recollection, particularly if strong familiarity prompts a person to conjure incorrect details that feel like accurate recollection. Thus, although this "Remember/Know" approach (as it is called in the field) is an outdated method that cannot be counted on to accurately reveal whether judgments are based on accurate recollection as opposed to simply based on the level of familiarity elicited by the stimulus, it can allow us to examine 1) how often illusory recollection occurred (i.e., the rate of reported subjective recall in the unstudied condition) and 2) if participants' subjective impressions are consistent with the hypothesis that stronger perceived familiarity with a novel face is more likely to be accompanied by a subjective sense of recollection.

The Probability of Reporting a Subjective Sense of Recall

A repeated-measures ANOVA was performed for the proportion of "Yes" responses to the subjective recall prompt (see Table 3 below for proportions of *Yes* responses) in each of the feature conditions, revealing a main effect of Feature Condition, $F_{subjects}(3, 366) = 123.89$, $MSE = 1.68$, $p < .001$, $\eta_p^2 = .503$, $BF_{10} = 9.87 \times 10^{51}$; $F_{items}(3, 237) = 79.67$, $MSE = 1.07$, $p < .001$, $\eta_p^2 = .502$, $BF_{10} = 7.75 \times 10^{32}$.[1] Participants more often said *yes,* that they subjectively recall previously seeing the face, for the faces that were studied with a mask than unstudied, $t_{subjects}(120) = -8.40$, $SE = .09$, $p < .001$, $d = -1.16$, $BF_{10} = 1.09 \times 10^{17}$; $t_{items}(79) = -11.02$, $SE = .12$, $p < .001$, $d = -1.20$, $BF_{10} = 9.93 \times 10^{13}$. Participants also responded *yes* more often to faces studied with sunglasses than unstudied, $t_{subjects}(120) = -6.74$ $SE = .11$, $p < .001$, $d = -1.19$, $BF_{10} = 1.52 \times 10^{18}$; $t_{items}(79) = -11.02$, $SE = .12$, $p < .001$, $d = -1.24$, $BF_{10} = 9.93 \times 10^{13}$. There was no significant difference in the proportion of *yes* responses for the mask and sunglasses feature conditions, $t_{subjects}(120) = .60$, $SE = .08$, $p = .551$, $d = -.06$, $BF_{01} = 1.61$; $t_{items}(79) = .21$, $SE = .17$, $p = .838$, $d = .00$, $BF_{01} = 8.11$. Participants also responded *yes* more often to faces from the mask + sunglasses condition than to faces from the mask-only condition, $t_{subjects}(120) = -4.19$, $SE = .07$, $p < .001$, $d = -.67$, $BF_{10} = 4.78 \times 10^{10}$; $t_{items}(79) = -7.4$, $SE = .10$, $p < .001$, $d = -.60$, $BF_{10} = 325994.01$, and than to faces from the sunglasses-only condition, $t_{subjects}(120) = -4.33$ $SE = .08$, $p < .001$, $d = -.58$, $BF_{10} = 2.23 \times 10^{8}$; $t_{items}(79) = -5.48$, $SE = .14$, $p < .001$, $d = -.62$, $BF_{10} = 6763.98$. A one-sample *t*-test, comparing the proportion of subjective recall for the unstudied condition to zero (since they objectively cannot recall an unstudied face), revealed that the proportion of *yes* responses to unstudied faces was significantly greater than zero, $t_{subjects}(122) = 25.32$, $SE = .37$, $p < .001$, $BF_{10} = 8.35 \times 10^{46}$; $t_{items}(79) = 22.65$, $SE = .02$, $p < .001$, $BF_{10} = 8.43 \times 10^{32}$.

Table 3. Proportion of feature conditions judged as *Yes* to the recall prompt for subjects and items.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	.37	.16	.37	.15
Mask	.55	.15	.55	.15
Sunglasses	.56	.16	.55	.14
Mask + Sunglasses	.65	.15	.64	.15

Familiarity Ratings as a Function of Subjective Sense of Recall vs. Non-recall

A 2 Subjective Recall Status (*subjectively recalled vs. subjectively unrecalled*) x 4 Feature Condition (*unstudied vs. mask-only vs. sunglasses-only vs. mask + sunglasses*) repeated-measures ANOVA revealed that there was a main effect of Feature Condition on the familiarity ratings given to whole faces at test, $F_{subjects}(3,345) = 45.19$, $MSE = 18.45$, $p < .001$, $\eta_p^2 = .277$, $BF_{10} = .07$; $F_{items}(3, 237) = 29.12$, MSE = 15.05, $p < .001$, $\eta_p^2 = .269$, $BF_{10} = > 150$.[1] There was also a main effect of Subjective Recall Status such that, overall, participants' subjective indications of recall success vs. failure separated into higher vs. lower familiarity ratings, respectively, $F_{subjects}(1, 118) = 1073.58$, $MSE = 6562.53$, $p < .001$, $\eta_p^2 = .901$, $BF_{10} = > 150$; $F_{items}(1, 79) = 5438.51$, $MSE = 4306.03$, $p < .001$, $\eta_p^2 = .986$, $BF_{10} = > 150$. There was also a significant interaction, $F_{subjects}(3, 354) = 16.55$, $MSE = .39$, $p < .001$, $\eta_p^2 = .123$; $F_{items}(3, 237) = 12.38$, $MSE = 4.69$, $p < .001$, $\eta_p^2 = .135$, which was such that the effects of the Feature Condition on familiarity ratings were larger when participants subjectively though they were experiencing

[1] There were 2 subjects who were removed from the analyses for the subjective sense of recall, as they did not respond "yes" to any of the faces for one or more of the feature conditions. For the no sense of recall analyses, 2 subjects were removed due to also not meeting the inclusive criterion of responding "no" to at least 1 face for each of the feature conditions.

recall success than when they subjectively thought that they were not. For BF analyses, due to the high number for evidence in favor of the Subjective Recall Status main effect (> 150), it exceeded the precision of JASP. Thus, the program was unable to determine the evidence in favor of the interaction for the subjects. For the items analysis, this error also occurred, as both the Subjective Recall Status and Feature Condition main effects were > 150.

Familiarity Ratings During the Subjective Sense of Recall

A repeated-measures ANOVA performed on the average familiarity ratings for each face condition (unstudied, mask-only, sunglasses-only, mask + sunglasses) when participants subjectively perceived that they did recall it (see table 4 below), revealed a significant main effect $F_{subjects}(3, 360) = 53.05$, $MSE = 23.46$, $p < .001$, $\eta_p^2 = .307$, $BF_{10} = 1.66 \times 10^{25}$; $F_{items}(3, 237) = 44.07$, $MSE = 17.71$, $p < .001$, $\eta_p^2 = .358$, $BF_{10} = 1.24 \times 10^{20}$.[1] Overall, when participants had the subjective impression that they recalled previously studying the face, higher familiarity ratings were given for whole test faces that had been studied wearing a mask compared to the unstudied faces, $t_{subjects}(120) = -8.40$, $SE = .09$, $p < .001$, $d = -.60$, $BF_{10} = 6.53 \times 10^{10}$; $t_{items}(79) = -7.18$, $SE = .11$, $p < .001$, $d = -1.08$, $BF_{10} = 3.04 \times 10^{7}$. Likewise, they gave higher familiarity ratings to whole test faces that had appeared at study wearing a pair of sunglasses compared to the unstudied faces, $t_{subjects}(120) = -6.74$, $SE = .11$, $p < .001$, $d = -.56$, $BF_{10} = 1.56 \times 10^{7}$; $t_{items}(79) = -7.44$, $SE = .10$, $p < .001$, $d = -.97$, $BF_{01} = 8.96 \times 10^{7}$. When comparing the whole test faces that were studied with a mask to faces studied with a pair of sunglasses as a function of subjectively recalling the face as previously studied, there was no significant difference, $t_{subjects}(120) = .60$, $SE = .08$, $p = .551$, $d = .04$, $BF_{01} = 8.33$; $t_{items}(79) = .49$, $SE = .11$, $p = .627$, $d = .06$, $BF_{01} = 7.22$. This suggests that the increase in familiarity-detection among the different exposed feature segments of the face was roughly equal, consistent with what was found by

Carlaw et al. (2022). Finally, participants gave higher familiarity ratings to the whole test faces that had appeared at study in two separate instances, once wearing a mask and another with sunglasses (mask + sunglasses) than to the test faces that had only appeared earlier with a mask, $t_{subjects}(120) = -4.19$, $SE = .07$, $p < .001$, $d = -.25$, $BF_{10} = 324.25$; $t_{items}(79) = -4.56$, $SE = .07$, $p < .001$, $d = -.47$, $BF_{10} = 943.60$, or only with a pair of sunglasses, $t_{subjects}(120) = -4.33$, $SE = .08$, $p < .001$, $d = -.29$, $BF_{10} = 526.20$; $t_{items}(79) = -3.76$, $SE = .10$, $p < .001$, $d = -.51$, $BF_{01} = 68.81$.

Table 4. Average familiarity ratings for each type of feature studied when the participant responded *Yes* to the subjective recall prompt for subjects and items.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	6.85	1.30	6.69	.76
Mask	7.60	1.17	7.49	.72
Sunglasses	7.55	1.21	7.44	.79
Mask + Sunglasses	7.88	1.03	7.81	.65

As for the level of increase in perceived familiarity ratings during subjective impressions of recall (see Table 5 below), a one-sample *t*-test was run for the average level of increase for mask + sunglasses condition ($M = 1.03$) against 1.46 (i.e., the combined average increase for the mask and sunglasses conditions) which revealed there was a significant difference, $t_{subjects}(120) = -4.59$, $SE = .09$, $p < .001$, $BF_{10} = 1308.28$. For the items, the one-sample *t*-test for mask + sunglasses condition ($M = 1.11$) was run against 1.53 which revealed there was a significant difference, $t_{items}(79) = -3.79$, $SE = .11$, $p < .001$, $BF_{10} = 73.10$. As with the overall pattern above, the level of additional perceived familiarity increase afforded by the co-occurring combination of separately familiarized facial feature sets within the whole face (over and above the individual feature set familiarization) was sub-additive.

Table 5. Mean level of familiarity-rating *increase* for each type of feature studied when the participant responded *Yes* to the subjective recall prompt for subjects and items.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	-	-	-	-
Mask	.75	.99	.79	.99
Sunglasses	.71	1.15	.74	.89
Mask + Sunglasses	1.04	1.02	1.11	.99

Familiarity Ratings During the Subjective Sense of non-Recall

A repeated-measures ANOVA performed on the average familiarity ratings for each face condition (unstudied, mask, sunglasses, mask + sunglasses) when participants subjectively perceived that they did *not* recall (see Table 6 below), revealed a significant main effect $F_{subjects}(3,360) = 4.61$, $MSE = 1.62$, $p = .004$, $\eta_p^2 = .037$, $BF_{10} = 3.97$; $F_{items}(3, 237) = 4.10$, $MSE = 2.03$, $p = .007$, $\eta_p^2 = .049$, $BF_{10} = 3.43$.[1] Although the means trended toward higher familiarity ratings among whole test faces that had been studied wearing a mask compared to the unstudied faces, this pattern only approached significance, $t_{subjects}(120) = -1.84$, $SE = .08$, $p = .069$, $d = -.11$, $BF_{01} = 1.96$; $t_{items}(79) = -1.64$, $SE = .08$, $p = .106$, $d = -.21$, $BF_{01} = 2.26$. This pattern reached significance when comparing faces that had been studied with sunglasses compared to unstudied faces, $t_{subjects}(120) = -2.37$, $SE = .07$, $p = .020$, $d = -.12$, $BF_{10} = 1.43$, but only in the subjects analysis; it was not significant in the items analysis, $t_{items}(79) = -1.22$, $SE = .08$, $p = .226$, $d = -.15$, $BF_{01} = 3.94$. There was no significant difference between whole test faces that had been studied with a mask and whole test faces that had been studied with sunglasses, $t_{subjects}(120) = -.23$, $SE = .07$, $p = .817$, $d = -.01$, $BF_{01} = 9.66$; $t_{items}(79) = .36$, $SE = .10$, $p = .723$, $d = .06$, $BF_{01} = 7.64$. Finally, although the pattern trended toward participants giving higher familiarity ratings to the whole test faces that had appeared at study in two separate instances, once wearing a mask and another with sunglasses (mask + sunglasses) than to the test faces that had only appeared earlier with a mask, these patterns did not reach significance, $t_{subjects}(120) = -1.97$, $SE = .07$, $p = .052$, $d = -.10$, $BF_{01} = 1.52$; $t_{items}(79) = -1.83$, $SE = .13$, $p = .072$, $d = -.27$, $BF_{01} = 1.67$. When comparing the familiarity ratings for the mask + sunglasses studied faces than for the faces that were only studied wearing a pair of sunglasses, the pattern was not significant in the subjects analysis, $t_{subjects}(120) = -1.44$, $SE = .09$, $p = .152$, $d = -.10$, $BF_{01} = 3.56$, but was significant in the items analysis, $t_{items}(79) = -2.14$, $SE = .13$, $p = .035$, $d = -.32$, $BF_{10} = 1.07$.

Table 6. Average familiarity ratings for each type of feature studied when the participant responded *No* to the subjective recall prompt.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	2.05	1.30	2.02	.46
Mask	2.19	1.36	2.15	.71
Sunglasses	2.20	1.31	2.11	.70
Mask + Sunglasses	2.33	1.39	2.39	1.03

As for the level of increase in familiarity ratings for *No* subjective recall, that were based on the type of feature studied (see Table 7 below), a one-sample *t*-test was run for the average level of increase for mask + sunglasses condition ($M = .28$) against .29 (i.e., the added together average increase for the mask and sunglasses conditions) which revealed there was no significant difference, $t_{subjects}(120) = -.09$, $SE = .09$, $p = .932$, $BF_{01} = 9.87$. For the items, the one-sample *t*-test for ($M = .37$) was run against .23 which revealed there was no significant difference, $t_{items}(79) = 1.14$, $SE = .13$, $p = .256$, $BF_{01} = 4.34$.

Table 7. Mean level of familiarity-rating increase for each type of feature studied when the participant responded *No* to the subjective recall prompt.

	Subjects		Items	
Feature Condition	*M*	*SD*	*M*	*SD*
Unstudied	-	-	-	-
Mask	.14	.82	.13	.74
Sunglasses	.15	.71	.10	.71
Mask + Sunglasses	.28	.93	.37	1.13

Chapter 4 – Discussion

Overview of the Present Study

The current experiment evaluated if familiarizing parts of human faces (through partial occlusion from surgical masks or sunglasses) would increase the perceived familiarity of the later whole face in which those parts were embedded. The experiment also examined if separately familiarizing different parts of a face at different points in time (e.g., viewing a surgical-mask-covered face in one instance and then the viewing same face wearing only sunglasses in another instance) would work to combinatorically increase the later perceived familiarity with the whole face in which those familiarized parts are embedded. Past research with a different type of stimulus—music—suggests that familiarizing different isolated feature types (rhythm and pitch) leads to increased perceived familiarity with a whole musical piece in which those features are embedded. In addition, separately familiarizing each of these isolated feature segments from a musical piece combinatorically boosts the later perceived familiarity with the whole in which those features are embedded relative to when only one of the feature types (rhythm or pitch) had been familiarized (McNeely-White et al., 2021).

However, unlike musical stimuli, facial stimuli are thought to be processed more holistically although they can be broken-down into features (Cabeza & Kato, 2000; Tanaka & Simonyi, 2016). Some have also suggested that faces may only be segmented into features when the face contains an occlusion over some of its parts. Accordingly, if a whole face is presented unoccluded, a holistic face processing system will automatically "kick in" whereas, if a face contains an occlusion over some of its parts, a feature-based processing system will take over

(Manley et al., 2019). If so, one would expect the feature-based processing system to dominate during encoding of occluded faces in the present study while the holistic processing system dominates at test when whole faces are presented. This might be expected to lead to *no increase* in familiarity with a whole face at test with the feature familiarization manipulation. This was examined in the present study, to determine if holistic processing has a role in the familiarity-detection of faces—if this were the case, then one would see such a pattern:

(mask + sunglasses) = [(mask-only) = (sunglasses-only)] = unstudied faces

Contrary to this idea, however, the results of the present study show (for both subjects and items analyses) that for overall recognition, facial component part familiarization *did* boost the overall familiarity of the whole face in which they were later embedded, and for both the nose and mouth features (sunglasses condition) and the eyes/eyebrow features (surgical masks condition). Moreover, separate exposures to each feature type served to increase the perceived familiarity for the whole face at test over and above the level of familiarity boost received for either feature type alone. Furthermore, there were no significant differences among the mask-only and sunglasses-only conditions (1x exposure), consistent with familiarity-detection being equal among parts of the face (Carlaw et al., 2022). Thus, faces may be processed as segmented features and boost the overall familiarity of the face, suggesting that holistic processing did not obligatorily "kick in" to replace parts processing when the whole face was available.

Theoretical Implications

The Segmenting of Whole Faces into Component Features

The overall familiarity ratings pattern demonstrates that different familiarized feature types combine to increase the perceived familiarity with a whole face. This overall pattern suggests that whole faces *are* broken-down into component features and are not processed solely, or even primarily, holistically. This is theoretically informative because although past research had suggested that familiarity-detection with faces could be based on component features of the face, as familiarity-detection could occur for unidentifiable faces that were occluded by either a surgical mask or sunglasses (Carlaw et al., 2022), it had remained unclear if faces presented as wholes were actually segmented into component features or parts by the mind (e.g., Manley et al., 2019).

Prior to the experiment, it was unclear if facial features separately stored in memory would be able to combine to systematically increase the perceived familiarity for a whole otherwise novel face (that had never before been seen in its entirety). If the results were to have shown no differences among feature conditions, this would have pointed toward the idea that recognizing whole faces relies on a holistic face processing system. However, finding such a systematic increase is consistent with results that have been shown with other types of stimuli (e.g., words, music) that have suggested that separately encoded stimulus features combine across memory traces to boost the familiarity of a later whole in which they are embedded (Huebert et al., 2022; McNeely-White et al., 2021, Ryals & Cleary, 2012).

In further examining how the separately encoded facial feature sets (i.e., nose-mouth vs. eyes/eyebrows) in the present study combined across instances in memory to produce an increased sense of familiarity with the whole test face in which they were embedded, we found that the boost afforded by the co-occurring combination of previously separately encoded feature sets was sub-additive in nature. That is, although their combined presence in the whole face boosted familiarity more than the presence of either feature set alone, the level of boost was less than what would be expected if the increase afforded by either feature type alone had summed together. Thus, the current results show that, contrary to being either equal to or greater than the sum of its parts, the perceived familiarity of the whole face is *less than* the sum of its parts:

(masks + sunglasses) < [(mask-only) + (sunglasses-only)] > unstudied faces

As for why this pattern occurred, one possible reason could be if the variances of the distributions of the Feature Conditions are unequal. McNeely-White et al. (2021) and McNeely-White et al. (2022) showed that separately encoded feature sets tend to show an additive combination in the level of boost to perceived familiarity intensity in situations where the signal and noise distributions (or various conditions' distributions) are roughly equal. One factor that affected whether signal and noise distributions were equal is the extent to which existing knowledge for the stimuli is already in the memory base. The more novel the stimuli, the less equal the variances tend to be with increasing experimental familiarization. Thus, because the facial stimuli used in the present study were all novel, no existing familiarity distribution for the faces would be expected to have existed in participants' knowledge-base prior to the experiment, making it highly likely that the variances would have been unequal across the conditions (e.g., Hintzman, 1988; McNeely-White et al., 2021; McNeely-White et al., 2022), and unequal variances can contribute to patterns of sub-additive familiarity increase levels (e.g., McNeely-White et al., 2022). Although this was not examined in the present study, as it was not the focus, it is a worthwhile question for future research to explore as a possible reason for the pattern.

Participants' Subjective Impressions of Recall

Participants' subjective impressions of recall revealed some striking patterns. The first concerns the rates of reported subjective impressions of recall: Participants exhibited a relatively high rate of *false* impressions of recall, as shown by the fact that the probability of giving a "Yes" response to the subjective recall prompt was .37 in the unstudied condition. This pattern, taken together with prior research suggesting that false subjective recollective experience increases with increased familiarity-detection (Huebert et al., 2022) and that participants make more commission errors with increased familiarity-detection (Carlaw et al., 2022), points toward the possibility that many of the reports of subjective recall may have been illusory and familiarity-driven.

Consistent with this idea is the second striking finding pertaining to participants' subjective impressions of recall, which concerns how high the mean familiarity ratings were for unstudied faces during subjective impressions of recall. These were 6.85 across subjects and 6.69 across items, which are substantially higher than any of the familiarity ratings given during impressions of non-recall in any of the Feature Conditions. Overall, this pattern of high familiarity ratings given among instances of subjective impressions of illusory recall in the unstudied condition is consistent with the idea that strong sensations of familiarity may have been driving the impressions of illusory subjective recall. Taken together with prior findings pointing toward familiarity-driven illusory recollective experience, whereby systematically increasing cue familiarity drives corresponding increases in illusory recollection (Huebert et al., 2022). This pattern also points toward the possibility that the systematic increases in familiarity ratings shown across the Feature Conditions during subjective impressions of recall may have themselves been driving the subjective impressions of recall. In fact, it is possible that the increasing perceptions of familiarity across the Feature Conditions underlie the increased reports of impressions of recall across those conditions (Table 3).

The present patterns (that also suggest that perceived familiarity may be driving the illusory impressions of recall) are also quite remarkable considering how impairing facial occlusions like surgical masks are known to be (e.g., Carlaw et al., 2022; Carragher & Hancock, 2020; Freud et al., 2020). Specifically, surgical masks and sunglasses have been widely demonstrated to severely impair face processing and identification. In Carlaw et al.'s study, for example, the presence of a surgical mask impaired identification of a known, would-be identifiable celebrity by roughly 2/3 (only about 1/3 were identifiable), and the presence of sunglasses impaired such identification by roughly 3/4 (only about 1/4 were identifiable). Considering that all occluded faces presented in the encoding phase of the present study were never before seen by participants, it is quite striking how often they claimed to be recalling the face from the encoding phase when presented with the whole face containing a studied feature set. As shown in Table 3, participants in the present study claimed to recall the studied face more than half of the time across all of the studied feature conditions. This striking contrast to past research on mask- or sunglasses-occluded faces raises the fascinating possibility that participants in the present study were experiencing familiarity-driven recollective confabulation (Huebert et al., 2022).

Limitations and Future Directions

One limitation of the present study concerns a key difference between the facial feature-types separately encoded in the present study and the musical feature-types separately encoded in McNeely-White et al. (2021)'s study. Specifically, whereas the rhythm and pitch of musical pieces are isolable from one another in non-overlapping ways, the facial feature types used in the present study may have still had overlapping features with one another, such as the particular skin pigment of the face. Thus, features that were not of interest in the present study may have been inadvertently repeated across encoding instances in the mask + sunglasses condition. Given that repetition has previously been used to increase familiarity (Rhodes et al., 2008), it is possible that the two presentations of parts of one face (mask + sunglasses condition) may have increased the overall familiarity for the whole face due to repetition of an overlapping feature like skin pigment occurring in both instances. Future research should aim to address this in some manner, such as by holding skin pigment constant across the face part encoding instances, or by directly examining whether the present patterns can be shown for skin pigments themselves, as opposed to other facial features.

The aforementioned discussion also highlights a key difference between the present study and McNeely-White et al.'s (2021) study of familiarizing isolated music features. Namely, facial features may be more difficult to disentangle from other facial feature types than musical features are from one another. Another notable key difference between the two studies is that we utilized a rating scale for measuring familiarity, but McNeely-White et al. used a yes-no familiarity judgment. Thus, the comparisons that McNeely-White et al. were of the probability of saying *Yes* familiar rather than an average familiarity rating. Furthermore, the musical stimuli for McNeely-White et al.'s study were of well-known piano songs which allowed for observing the probability of saying *Yes* familiar as a function of identification status. Because all the faces used in the present study were novel, this was not possible. Future research should utilize celebrity face stimuli as a closer face analogy to McNeely-White et al.'s music study. Doing so would additionally allow for investigation of the yet unknown question of whether a holistic processing mechanism would only obligatorily "kick in" if a whole face has been processed as a whole previously in the participant's experience (as a celebrity face would). Given these key differences across studies, the current findings should not serve as a direct comparison between the different types of features (music and faces), but rather as a precedent for understanding differences in familiarity among different stimulus types and feature-types.

It is recognized that the Remember/Know method is considered to be an antiquated and unreliable method of trying to separate out the contributions of recollection versus familiarity (e.g., Rotello, et al., 2005; Wais et al., 2008). Recent research also suggests that strong familiarity might subjectively be *misinterpreted* as recollection—particularly if strong familiarity prompts a person to conjure incorrect details that feel like accurate recollection (Carlaw et al., 2022; Huebert et al., 2022). Thus, future research should examine whether the present pattern holds up in situations where face identifiability or recollection can be objectively examined to separate out objective instances of recall success vs. failure. This would include the reverse of the method used by Carlaw et al. (2022), whereby occluded would-be identifiable celebrity faces are presented in the encoding phase and the whole identifiable celebrity faces are presented at test.

The current experiment used novel faces to first establish any systematic increase in perceived familiarity among whole faces never seen before in their entirety. Accordingly, the present results inform theory regarding the supposition that whole faces might obligatorily trigger a holistic face processing system (Manley et al., 2019), but leave open the question of the extent to which familiarity-driven recollective confabulation or familiarity-driven false recollective experience may have been occurring. A follow-up to the present study using known celebrity faces instead of novel faces could elucidate this issue by providing a way of objectively indexing the accuracy of participants' recollective responses (because participants would supply identification responses or partial responses that can be objectively assessed for accuracy).

Another limitation of the present study is that it cannot yet be fully ascertained to what degree the low familiarity ratings given during subjective impressions of non-recall were driving the subjective impressions of non-recall in the first place. If high familiarity can drive illusory recollective experience, as suggested above, then it is possible that low familiarity drove the subjective impressions of non-recall and also the possible mixed patterns pertaining to the Feature Conditions across these instances. This is another reason why it is important for future research to use famous faces in the present paradigm to examine whether the same pattern is found when famous faces and objective indices are used.

Familiarity intensity appears to be connected to curiosity, which creates a drive to seek the information. Metcalfe and colleagues (2017) found that when participants were in a tip-of-the-tongue state (i.e., the feeling that one knows the answer to a question and is on the verge of retrieving it, but cannot access it) they were more likely to want to see the answer to the question and more likely to make a commission error in a failed attempt to come to the answer on their own. Since high familiarity appears to drive illusory recollection, a follow-up experiment objectively separating recollection and familiarity for celebrity faces should also examine commission and omission errors. Determining of increased perceived familiarity intensity drives up commission errors and/or generation of confabulatory information to explain the familiarity would help to further determine if illusory recollection was what may have been driving the high rates of subjective recall found in the current study.

If future research were found to show that high familiarity drives illusory recollection for a whole face when only parts were studied, this may have implications for understanding eyewitness lineup accuracy. It may be that high familiarity driving illusory recollection is associated with false alarms in a lineup (falsely identifying a suspect when the perpetrator is absent from the lineup) for instances of when the perpetrator wore a face occlusion. Thus, familiarity should be explored as a potential measure for eyewitness lineup procedures.

One limitation of the current study and face processing research in general is that there is uncertainty of what constitutes a facial feature. Diego-Mas (2020) raised this as a point of concern in the face processing domain, as it is unclear how facial features are defined. Diego-Mas provided an example, that the eyes can be broken-down into even smaller features such as the upper and lower eyelids, so it is unclear exactly what a "feature" is defined as for the face. Is it the smaller features that make up the eyes (e.g., upper and lower eyelids, etc.) or is it the entirety of small features that make of the eyes-feature? Or could one argue that the section of the face that is available is a feature? For example, in the current study, the faces that were occluded by a mask covering the nose, mouth, and chin—but does that serve as one "feature", is it those ladder features, or is it the smaller features (e.g., cupid's brow, nostrils, shapes of the nose, etc.). Thus, it is unclear how faces should be broken into features for study. Future research should consider this by trying to isolate features as strictly as possible by being more precise with the covering of the features and examine those "smaller" features in separate study. It may be also a point of interest to isolate more "individual" features such as separate the nose and mouth to determine if familiarity detection is equal for these separate features compared to the eyes, as opposed to sections of the face serving as features. Another set of features that could be potentially isolated are the eyes and eyebrows, as these could be defined as separate features. Thus, future studies should examine different variations of how the face can be broken into features—as it is possible that the face composes of different sizes of features (e.g., the smaller features make up the larger features, which make up the whole face).

Future research should also objectively separate features and configuration, as the current findings may cast doubt on the importance of configural information in processing faces, as this seemed to not play a role in boosting the overall familiarity multiplicatively. If the faces were studied twice (mask + sunglasses condition), then it would likely be that having additional configural information would boost the familiarity for the face more than in an additive fashion. Researchers have argued that configural information is not relied on for recognition of familiar faces (Burton et al., 2015; Sandford & Burton, 2014). However, the current experiment did not directly examine features versus configuration. This should be considered in future studies to determine if familiarity-detection of novel and celebrity faces is based on features or that configuration is included. This could be examined by using a variation of Sandford and Burton's (2014) method, by stretching the face images (changing the configuration) and examining if the configuration is changed, then will this prompt familiarity for that face in the absence of recall. If so, then this would be more determinant of whether facial features alone are what drives recognition, which would be consistent with other types of stimuli that have been studied.

References

Brown, A. S. (2020). The butcher on the bus experience. In A. M. Cleary & B. L. Schwartz (Eds.), *Memory quirks: The study of odd phenomena in memory* (pp. 224–247). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429264498-17

Bruyer, R. (2011). Configural face processing: A meta-analytic survey. *Perception*, *40*(12). 1478–1490. https://doi.org/10.1068/p6928

Burton, A. M., Schweinberger, S. R., Jenkins, R., Kaufmann, J. M. (2015). Arguments Against a Configural Processing Account of Familiar Face Recognition. *Perspectives on Psychological Science, 10*(4), 482–496. https://doi.org/10.1177/1745691615583129

Cabeza, R., & Kato, T. (2000). Features are also important: contributions of featural and configural processing to face recognition. *Psychological science*, *11*(5), 429–433. https://doi.org/10.1111/1467-9280.00283

Carlaw, B. N., Huebert, A. M., McNeely-White, K. L., Rhodes., M. G., & Cleary, A. M. (2022) Detecting a familiar person behind the surgical mask: Recognition without identification among masked vs. sunglasses-covered faces. *Cognitive Research: Principles and Implications*, *7*(1), 90. https://doi.org/10.1186/s41235-022-00440-3

Carragher, D. J., & Hancock, P. (2020). Surgical face masks impair human face matching performance for familiar and unfamiliar faces. *Cognitive Research: Principles and Implications*, *5*, 59, 1–15. https://doi.org/10.1186/s41235-020-00258-x

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*(1), 37–60. https://doi.org/10.3758/BF03210740

Cleary, A. M. (2002). Recognition with and without identification: Dissociative effects of meaningful encoding. *Memory & Cognition, 30,* 758–767.

Cleary, A. M. (2002). Recognition with and without identification: Dissociative effects of meaningful encoding. *Memory & Cognition*, *30*, 758–767. https://doi.org/10.3758/BF03196431

Cleary, A. M. (2019). The biasing nature of the tip-of-the-tongue experience: When decisions bask in the glow of the tip-of-the-tongue state. *Journal of Experimental Psychology: General*, *148*(7), 1178–1191. https://doi.org/10.1037/xge0000520

Cleary, A. M. & Specker, L. E. (2007). Recognition without face identification. *Memory & Cognition, 35*(7), 1610–1619. https://doi.org/10.3758/BF03193495

Cleary, A. M., & Greene, R. L. (2000). Recognition without identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1063–1069. https://doi.org/10.1037/0278-7393.26.4.1063

Cleary, A. M., & Greene, R. L. (2001). Memory for unidentified items: Evidence for the use of letter information in familiarity processes. *Memory & Cognition, 29*(3), 540–545. https://doi.org/10.3758/BF03196405

Cleary, A. M., & Greene, R. L. (2004). True and false memory in the absence of perceptual identification. *Memory, 12*(2), 231-236. https://doi.org/10.1080/09658210244000577

Cleary, A. M., & Greene, R. L. (2005). Recognition without perceptual identification: A measure of familiarity? *Quarterly Journal of Experimental Psychology, 58A*(6)*,* 1143–1152. https://doi.org/10.1080/02724980443000665

Cleary, A. M., Konkel, K. E., Nomi, J. N., & McCabe, D. P. (2010). Odor recognition without identification. *Memory & Cognition, 38,* 452–460. https://doi.org/10.3758/MC.38.4.452

Cleary, A. M., Langley, M. M., & Seiler, K. R. (2004). Recognition without picture identification: Geons as components of the pictorial memory trace. *Psychonomic Bulletin & Review, 11*(5), 903–908. https://doi.org/10.3758/BF03196719

Cleary, A. M., Ryals, A. J., & Nomi, J. S. (2013). Intuitively detecting what is hidden within a visual mask: Familiar-novel discrimination and threat detection for unidentified stimuli. *Memory & Cognition, 41*(7)*,* 989–999. https://doi.org/10.3758/s13421-013-0319-4

Cleary, A. M., Ryals, A. J., & Wagner, S. R. (2016). Recognition during recall failure: Semantic feature matching as a mechanism for recognition of semantic cues when recall fails. *Memory & Cognition*, *44*(1), 50–62. https://doi.org/10.3758/s13421-015-0545-z

Cleary, A. M., Winfield, M. M., & Kostic, B. (2007). Auditory recognition without identification. *Memory & Cognition, 35*(8)*,* 1869–1877. https://doi.org/10.3758/bf03192921

DeBruine, L. M. & Jones, B. C. (2017). *Face Research Lab London Set*. figshare https://doi.org/10.6084/m9.figshare.5047666.v5

Diego-Mas, J. A., Fuentes-Hurtado, F., Naranjo, V., & Alcañiz, M. (2020). The influence of each facial feature on how we perceive and interpret human faces. *i-Perception*, *11*(5), 2041669520961123. https://doi.org/10.1177/2041669520961123

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, *6*(3), 534–539. https://doi.org/10.1037/1528-3542.6.3.534

Freud, E., Giammarino, D.D., Stajduhar, A., Rosenbaum, R. S., Avidan, G., & Ganel T. (2022). Recognition of Masked Faces in the Era of the Pandemic: No Improvement Despite Extensive Natural Exposure. *Psychological Science.* https://doi.org/10.1177/09567976221105459

Freud, E., Stajduhar, A., Rosenbaum, R. S., Avidan, G., & Ganel, T. (2020). The COVID-19 pandemic masks the way people perceive faces. *Scientific Reports*, *10*(1), 22344. https://doi.org/10.1038/s41598-020-78986-9

Galton, F. (1883). *Inquiries into human faculty and its development*. Macmillan. https://galton.org/books/human-faculty/text/galton-1883-human-faculty-v4.pdf

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, *54*(6), 917–924. https://doi.org/10.1037/0022-3514.54.6.917

Hintzman, D. L. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551. https://doi.org/10.1037/0033-295X.95.4.528

Huebert, A. M., & Cleary, A. M. (2022). Do first and last letters carry more weight in the mechanism behind word familiarity? *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02093-1

Huebert, A. M., McNeely-White, K. L., & Cleary, A. M. (2022). Can cue familiarity during recall failure prompt illusory recollective experience? *Memory & Cognition*. https://doi.org/10.3758/s13421-021-01248-0

Kostic, B. & Cleary, A. M. (2009). Song recognition without identification: When people cannot “name that tune” but can recognize it as familiar. *Journal of Experimental Psychology: General, 138*(1), 146–159. https://doi.org/10.1037/a0014584

Kwart, D. G., Foulsham, T., & Kingstone, A. (2012). Age and beauty are in the eye of the beholder. *Perception*, *41*(8), 925–938. https://doi.org/10.1068/p7136

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863.

Langley, M. M., Cleary, A. M., Woods, J. & Kostic, B. (2008). Picture recognition without picture identification: A method for assessing the role of perceptual information in familiarity-based picture recognition. *Acta Psychologica, 127*(1)*,* 103–113. https://doi.org/10.1016/j.actpsy.2007.03.001

MacLeod, C. M. (2020). The butcher on the bus: A note on familiarity without recollection. *History of Psychology*, *23*(4), 383-387. https://doi.org/10.1037/hop0000178

Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 207–225). Lawrence Erlbaum Associates, Inc.

Manley, K. D., Chan, J., & Wells, G. L. (2019). Do masked-face lineups facilitate eyewitness identification of a masked individual? *Journal of Experimental Psychology: Applied*, *25*(3), 396–409. https://doi.org/10.1037/xap0000195

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. http://doi.org/10.1016/S1364-6613(02)01903-4

McNeely-White, K. L., McNeely-White, D. G., & Cleary, A. M. (2021). Global matching in music familiarity: How musical features combine across memory traces to increase familiarity with the whole in which they are embedded. *Journal of Memory and Language, 118,* 104217. https://doi.org/10.1016/j.jml.2020.104217

McNeely-White, K. L., McNeely-White, D. G., Huebert, A. M., Carlaw, B. N., & Cleary, A. M. (2022). Specifying a relationship between semantic and episodic memory in the computation of a feature-based familiarity signal using MINERVA 2. *Memory & Cognition*, *50*(3), 527–545. https://doi.org/10.3758/s13421-021-01234-6

Metcalfe, J., Schwartz, B. L., & Bloom, P. A. (2017). The tip-of-the-tongue state and curiosity. *Cognitive Research: Principles and Implications*, *2*(1), 31. https://doi.org/10.1186/s41235-017-0065-4

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*(5), 553–566. https://doi.org/10.1068/p3339

Monin B. (2003). The warm glow heuristic: when liking leads to familiarity. *Journal of Personality and Social Psychology*, *85*(6), 1035–1048. https://doi.org/10.1037/0022-3514.85.6.1035

Morris, A. L., Cleary, A. M., & Still, M. L. (2008). The role of autonomic arousal in feelings of familiarity. *Consciousness and Cognition, 17*(4), 1378–1385. https://doi.org/10.1016/j.concog.2008.04.005

Nguyen, H. T., Isaacowitz, D. M., & Rubin, P. A. D. (2009). Age- and fatigue-related markers of human faces: An eye-tracking study. *Ophthalmology, 116*(2). 355–360. https://doi.org/10.1016/j.ophtha.2008.10.007

Pazhoohi, F., & Kingstone, A. (2022). Unattractive faces are more attractive when the bottom-half is masked, an effect that reverses when the top-half is concealed. *Cognitive*

Research: Principles and Implications, *7*(1), 6. https://doi.org/10.1186/s41235-022-00359-9

Peynircioğlu, Z. F. (1990). A feeling-of-recognition without identification. *Journal of Memory & Language, 29*(4), 493–500. https://doi.org/10.1016/0749-596X(90)90068-B

Pinkham, A. E., Griffin, M., Baron, R., Sasson, N. J., & Gur, R. C. (2010). The face in the crowd effect: Anger superiority when using real faces and multiple identities. *Emotion*, *10*(1), 141–146. https://doi.org/10.1037/a0017387

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, *21*(1), 89–102. https://doi.org/10.3758/BF03211168

Rhodes, M. G., Castel, A. D., & Jacoby, L. L. (2008). Associative recognition of face pairs by younger and older adults: The role of familiarity-based processing. *Psychology and Aging*, *23*(2), 239–249. https://doi.org/10.1037/0882-7974.23.2.239

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139–253. https://doi.org/10.1080/13506285.2013.772929

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, *12*(5), 865-873. https://doi.org/10.3758/BF03196778

Ryals, A. J., & Cleary, A. M. (2012). The recognition without cued recall phenomenon: Support for a feature-matching theory over a partial recollection account. *Journal of Memory and Language*, *66*(4), 747–762. https://doi.org/10.1016/j.jml.2012.01.002

Sandford, A., & Burton, A. M. (2014). Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. *Cognition*, *132*(3), 262–268. https://doi.org/10.1016/j.cognition.2014.04.005

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46A*(2), 225–245. http://doi.org/10.1080/14640749308401045

Tanaka, J. W., & Simonyi, D. (2016). The "parts and wholes" of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *69*(10), 1876–1889. http://doi.org/10.1080/17470218.2016.1146780

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*(4), 471–491. https://doi.org/10.1111/j.2044-8295.1988.tb02747.x

Valentine, T., Darling, S., & Donnelly, M. (2004). Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychonomic Bulletin & Review*, *11*(3), 482–487. https://doi.org/10.3758/BF03196599

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779-804. https://doi.org/10.3758/BF03194105

Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/know judgments probe degrees of recollection. *Journal of Cognitive Neuroscience*, *20*(3), 400–405. https://doi.org/10.1162/jocn.2008.20041

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, *36*(12), 1546–1557. https://doi.org/10.1037/0022-3514.36.12.1546

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. https://doi.org/10.1037/h0027474

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science*, *26*(3), 212–217.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–759. https://doi.org/10.1068/p160747

Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *NeuroImage, 21*, 789–800. http://doi.org/10.1016/j.neuroimage.2003.09.034