Does odour facilitate face learning? An Event-Related Potentials (ERP) study.

Rui Su

How to cite:

Su, Rui (2026) Does odour facilitate face learning? An Event-Related Potentials (ERP) study. Masters thesis, Durham University.

Abstract

Face learning is a core social ability, but the contribution of olfactory cues to this process remains poorly understood. Building on models of auditory-visual integration in person recognition, we investigated whether an analogous Olfactory Recognition Unit (ORU) facilitates face learning. We tested if pairing odours with faces during learning enhances subsequent recognition. In Experiment 1, participants learned two identities from videos, each paired with a pleasant odour (rose or lilac). At test, recognition of the identities from highly variable images was assessed during odour re-presentation, while event-related potentials (ERPs) were recorded (N250: 200–400 ms; SFE: 400–600 ms). Congruent trials referred to faces tested with the same odour as during learning; incongruent trials referred to faces tested with a different odour; control faces were novel identities not previously learned. Experiment 2 replicated the procedure with a larger sample of Caucasian participants, more distinctive odours (rose vs. sandalwood), and behavioural measures only. Across experiments, recognition was superior for learned versus control faces, confirming robust learning. In Experiment 1, this was accompanied by neural learning effects, but the N250 effect was specific to the incongruent condition and was not found for the congruent condition, which might suggest a unique modulatory role for congruent odour on early neural responses. However, there was no reliable difference in recognition or ERP amplitudes between congruent and incongruent conditions in Experiment 1 and 2. A trend toward higher recognition for congruent faces in Experiment 2 did not survive correction for multiple comparisons. Thus, we found no robust behavioural or neural evidence that odour congruency facilitates face learning. An interaction between ethnicity and learning condition was observed in the ERP data: Caucasian participants showed N250 and SFE learning effects, whereas Asian participants did not. This may reflect different encoding strategies for other-race faces. Overall, these results do not support the hypothesis of an Olfactory Recognition Unit. Future studies should use natural body odours, include a no-odour control condition, and use larger samples to determine whether olfactory cues can ever become functionally significant for person recognition.

Does odour facilitate face learning? An Event-Related Potentials (ERP) study.

Rui Su

<u>Department of Psychology, Durham University</u>

Master by Research Dissertation

Supervisor: Prof. Holger Wiese

April 2, 2026

Table of content

Statement of Copyright

Acknowledgements

First, I would like to express my great thanks to my supervisor, Professor Holger Wiese. He has been professional, dedicated, and patient, providing constructive feedback on my thesis. Additionally, I wish to extend my thanks to all the members of the laboratory for their tremendous support whenever I encountered difficulties. I would also like to extend my sincere appreciation to my friends whom I met in Durham for their selfless support during my challenging times. Last but not least, I would like to thank my parents. Without their unwavering support, I would not have been able to accomplish this.

Does odour facilitate face learning? An Event-Related Potentials (ERP) study.

Face recognition is a critical component of human social cognition, enabling individuals to navigate interpersonal interactions, form relationships, and identify others across diverse environments. Humans are remarkably adept at recognizing familiar faces with high speed and accuracy, even under suboptimal conditions such as changes in lighting, viewpoint, or expression (Young and Burton, 2017). In contrast, the recognition of unfamiliar faces is significantly more error-prone and effortful, with limited improvement even after extensive training (Bruce et al., 1999; White et al., 2014).

These distinctions are supported by neurocognitive evidence. Electrophysiological markers, such as the N250 component, have been shown to index the processing of face familiarity, with increased amplitudes observed for familiar versus unfamiliar faces (e.g. Tanaka et al., 2006; Kaufman et al., 2009). Activation in face-sensitive brain regions, including the fusiform face area (FFA), further suggests distinct neural pathways for encoding and recognizing familiar identities (Duchaine and Yovel, 2015).

Despite robust findings on visual mechanisms in face learning, much less is known about the role of multisensory integration in facilitating this process. In real-world settings, faces are rarely encountered in isolation; they are accompanied by contextual cues such as voice, touch, and odour. Multisensory integration has been shown to enhance perceptual learning and memory in other domains (Seitz et al., 2006), yet its influence on face learning remains underexplored.

This thesis addresses a critical and understudied question: Can olfactory cues support or enhance the learning of unfamiliar faces? Specifically, it investigates whether the presence of distinctive odours during the encoding phase of unfamiliar face learning influences subsequent recognition, using event-related potentials (ERPs) to measure the underlying neural mechanisms.

Familiar and Unfamiliar Face Recognition

It has frequently been argued that humans possess a unique expertise in face recognition, driven by the social necessity of processing countless faces throughout a lifetime. Consequently, as we encounter specific types of information, we become experts at recognizing them. For instance, the other-race effect—where we recognize and learn faces from our own ethnic group better than from other ethnic groups—highlights the significance of experience with a specific category of stimuli (Young et al., 2012; Wiese et al., 2014; Bar-Haim et al., 2006; Kelly et al., 2007; Hayward et al., 2017). However, although our ability to recognize familiar faces reaches expert standards, unfamiliar face recognition is substantially more error-prone.

According to the standard criteria for expertise, there are three critical factors: (1) experience is important, (2) recognition must be accurate, and (3) it must be automatic (Young and Burton, 2017). For both unfamiliar and familiar face recognition, experience is crucial. For unfamiliar faces, phenomena such as the other-race effect underline the importance of prior experience at the category level. Similarly, studies on infants have shown that they recognize human faces better than monkey faces (Pascalis et al., 2002, 2005), and increased exposure to monkey faces improves recognition performance (Pascalis et al., 2005), highlighting experience's role in unfamiliar face recognition. Therefore, experience seems crucial for unfamiliar face recognition. Notably, genetic factors might also influence our face recognition ability, potentially challenging this first criterion, which will be discussed later.

Experience is undoubtedly significant for familiar face recognition at the individual level as well—the more familiar we are with a person, the better we recognize them. For example, greater exposure duration (Reynolds & Pezdek, 1992) and higher frequency of encounters (Xue et al., 2010) enhance face recognition performance. Additionally, greater variability during face learning (e.g., different hairstyles, makeup, settings) improves recognition performance (Ritchie & Burton, 2017; Corpuz & Oriet, 2022). Thus, both the quantity and quality of experience are crucial for familiar face recognition. Consequently, the first criterion (experience) is fulfilled for both familiar and unfamiliar face recognition.

Nevertheless, despite meeting the first criterion, accuracy differs significantly between familiar and unfamiliar face recognition. Numerous studies reveal that unfamiliar face recognition is often error-prone and worse than expected. In Bruce et al.'s (1999) face matching task, participants decided whether a target face appeared in an image array, with changing expression and viewpoint in some conditions, with unlimited time. Yet, the error rate remained around 30%. Even in simpler tasks, performance was relatively poor. For example, Kemp, Towell, and Pike's (1997) experiment required supermarket workers to match real faces to ID card images, resulting in false rejections of 10% of genuine IDs and false acceptances of 64% of fake IDs. Unfamiliar face recognition is thus worse than expected, likely because it relies on direct superficial feature comparisons (such as hair style and face shape)—or pictorial coding—vulnerable to subtle image variations.

Moreover, performance in unfamiliar face recognition is not easily improved by training and appears genetically influenced. Even highly trained passport officers show around a 10% error rate when matching volunteers to passport images, comparable to less experienced participants (White et al., 2014). However, a small percentage of the population, known as super-recognisers, are consistently excellent at unfamiliar face recognition—a skill that is highly heritable (Wilmer et al., 2010; Zhu et al., 2010). Thus, unfamiliar face recognition is often surprisingly inaccurate, relatively unaffected by training, and significantly genetically determined, challenging the criteria for expertise.

In contrast, familiar face recognition demonstrates clear expertise. Humans recognize familiar faces quickly, accurately, and automatically. In Jenkins et al.'s (2011) sorting task, unfamiliar participants sorted 40 images of two identities into an average of nine categories, while familiar participants correctly sorted them into two identities. Furthermore, unlike unfamiliar face matching tasks, familiar face recognition remains accurate despite significant variability across images (Ritchie et al., 2015; Burton et al., 1999). ERP studies also provide evidence of automatic familiar face recognition, even when participants perform unrelated tasks (Wiese et al., 2019; Neumann et al., 2011; Neumann & Schweinberger, 2008).

In summary, humans are experts in recognizing familiar faces but not unfamiliar faces, despite intuitive expectations. Unfamiliar face recognition is generally poor, minimally improvable by training, and strongly genetically influenced. Familiar face recognition, however, is fast, accurate, and automatic. Given the substantial differences between these two types of recognition—and considering that all familiar faces begin as unfamiliar—it is essential to understand how unfamiliar faces eventually become familiar.

Face Recognition Models and Face Learning

To understand how face learning occurs, it is essential to first review existing models of face recognition. Bruce and Young's (1986) face recognition model provides a foundational framework for explaining this process. According to this model, unfamiliar faces are initially recognized based on pictorial codes—surface-level visual details that change easily with viewing angle, lighting, or expression—which explains why unfamiliar face recognition is typically error-prone and inconsistent.

With repeated exposure to faces across various viewing conditions, observers gradually begin to extract structural codes. These structural codes represent stable, invariant characteristics of a face (such as 3D shape and surface reflectance information) and are stored in memory within a Face Recognition Unit (FRU). Each known face has a corresponding FRU. When a new image of a familiar face is encountered, the structural code extracted from that image is compared with the stored structural codes in the FRUs. The more similar the extracted and stored structural codes are, the stronger the activation of that FRU, resulting in recognition once a specific threshold is reached. Following activation of an FRU, semantic information about the person—such as occupation, nationality, or biographical details—stored in Person Identity Nodes (PINs) becomes accessible.

Research suggests that humans rely on different facial features to recognize familiar versus unfamiliar faces (Bonner et al., 2003; Ellis et al., 1979; Young et al., 1985). When a face is unfamiliar, we tend to use external features (e.g., face shape and hairstyle), which are highly variable across images. But when a face becomes familiar, people tend to rely on internal features (e.g., eyes, nose, and mouth), which are more invariant. Therefore, the process of face learning refers to a shift from external to internal features. In other words, when a face is learned, people tend to rely more on invariant features rather than variant ones, which aligns with the idea of Bruce and Young (1986).

However, accurately forming FRUs may not rely solely on invariant features—a view often described as bottom-up, which predominantly depends on general perceptual representations derived from visual input (Kramer et al., 2018). In reality, when we learn a face, it naturally varies considerably across instances (Burton et al., 2016). These variations are often intuitively dismissed as “noise” that interferes with identification. Yet rather than hindering recognition, such variability may in fact provide diagnostic cues that support learning—a perspective more aligned with a top-down view (Kramer et al., 2018). As mentioned previously, better learning effects are achieved through exposure to highly variable images. Therefore, the question arises whether variant features are also relevant for building an FRU?

Idiosyncratic Variant Features

While invariant features are crucial, recent evidence suggests that variant features—such as expressions, lighting conditions, and viewpoint changes—also play an important role in face learning (Benson & Perrett, 1993; Burton et al., 2005; Burton et al., 2011; Jenkins & Burton, 2008). If variant features contribute significantly to recognition, then recognition performance should improve with increased exposure to images that show greater variability. Indeed, studies confirm that recognition accuracy improves under conditions of high variability compared to low variability (Ritchie & Burton, 2017; Corpuz & Oriet, 2022).

Recent computational models further underscore the importance of both invariant and variant features in achieving robust face recognition performance (Burton et al., 2016). A statistical method—Principal Component Analysis (PCA)—is applied to analyze the variation across images of the same person, revealing that the variation of one face is different from that of another. In other words, each face has its own idiosyncratic variations, which are stable across images of the same person but differ between faces of different identities.

The idiosyncratic variability of faces explains why we are experts at recognizing familiar faces but not unfamiliar ones. For unfamiliar faces, we do not know whether the variation between images is idiosyncratic or irrelevant for identification. In other words, we do not know whether the variability between unfamiliar face images is within or between identities, making us easily distracted by them. In contrast, familiar face recognition is not disrupted by irrelevant variant features and focuses only on identity-specific cues. Indeed, supported by a computational modelling study, robust familiar face recognition is best mimicked when the model learns idiosyncratic within-person variability and integrates this information into perceptual representations. This finding suggests that the collaboration of top-down and bottom-up mechanisms is crucial for achieving human-like performance (Kramer et al., 2018).

Therefore, based on this idea, the face learning process involves separately learning both invariant features and idiosyncratic variant features. But how is a stable FRU formed if both variant and invariant features are important for recognizing familiar faces?

Stable Face Representation

The idea of averaging provides a powerful explanation. Building on Bruce's (1999) concept of "stability from variation," learning a face involves accumulating evidence across multiple images of the same face to form a stable, comprehensive representation. Burton and colleagues propose that stable representations of familiar faces in FRUs are achieved by averaging across multiple variable instances. To support this idea, Burton et al. (2005) developed a method in which they controlled face shape and averaged the texture of the face to create an averaged face. They found that averaged faces improve recognition performance in both computational models (Jenkins & Burton, 2008) and human participants. This suggests that averaging explains how robust face representations are formed.

Therefore, based on this idea, the more images and experience we have with a particular face, the more our cognitive system will automatically average the perceived faces and form a stable representation stored in the FRU. When a novel image of a familiar face is shown, we compare the perceived image with the averaged face representation in the

FRU—the greater the similarity between them, the stronger the FRU activation, leading to recognition of the face as familiar.

Notably, the familiar face recognition is not purely bottom-up comparison, which proposed by Bruce and Young (1986) suggesting that familiar face recognition involves a direct comparison between the structural code extracted from the incoming image and the stored representation. Instead, familiar face recognition appears to operate through an integration of top-down and bottom-up mechanism: after learning both the invariant and idiosyncratic variant features of a face, the visual system constructs an averaged face representation that filters out irrelevant variability. Rather than engaging in a pixel-by-pixel comparison, recognition is guided by diagnostic features—those consistent across encounters (invariant) and uniquely characteristic of the individual (idiosyncratic variant)—allowing for accurate identification across a wide range of visual input (Jenkins & Burton, 2011; Kramer et al., 2018).

ERP Work and ERP Face Recognition Models

Substantial previous work combines the EEG method with face perception and investigates how Event-Related Potentials (ERPs) are related to face processing. Schweinberger and Neumann (2016) integrated ERP findings with the Bruce and Young (1986) model. ERP is derived from EEG by averaging time-locked electrical brain responses from multiple trials to isolate the signal from background noise (Blackwood and Muir, 1990). Based on their serial model, when a face is perceived, a sequence of positive and negative EEG voltage changes occurs across five important time windows: Early visual processing (a positive brain potential occurring at around 100 ms, called P100), Face detection (a negative potential at approximately 170 ms, called N170), Prototypicality (a positive potential at around 200 ms, called P200), Face recognition (N250), and Identity-specific information (N400). It is important to note that the serial model does not imply strictly sequential processing; these stages may overlap (Schweinberger and Burton, 2003).

P100, N170, and P200 reflect the perceptual stage of face processing. Among them, P100 occurs earliest after stimulus onset and reflects the perception of low-level visual characteristics (e.g., luminance and contrast). N170 is more face-specific; face stimuli elicit a more negative-going amplitude compared to other visual stimuli (Eimer, 2011; Rossion and Jacques, 2008). Moreover, N170 is believed to be related to the extraction of structural codes (Carbon et al., 2005; Jemel et al., 2003). This is followed by P200, which reflects the "distance-to-norm" (DTN) for the perceived face stimulus—essentially the similarity of the target face to an average face (Schweinberger and Neumann, 2016). Notably, as part of the perceptual stage, for N170 the large majority of studies do not find a difference between familiar and unfamiliar faces (Alzueta et al., 2019; Andrews et al., 2017). P200, however, overlaps with the later N250 time window, and familiarity effects are often observed from approximately 200ms onwards (For review, see Wiese et al., 2024)

Crucially, at the recognition stage, a robust and consistent ERP component reflecting face familiarity is the N250(r). This is defined as a negative potential that typically occurs around 200–220 ms and peaks between 250–350 ms after face presentation (Wiese et al., 2024). Prior studies have shown that brief repetition of a face prior to its test presentation elicits a more negative N250 component than in non-repetition conditions—a phenomenon known as the N250r effect (Begleiter et al., 1995; Herzmann et al., 2004). This effect is more pronounced for familiar faces than for unfamiliar ones. However, in unfamiliar face conditions, the repeated face images must be tightly controlled, indicating high image-dependence (Zimmermann & Eimer, 2013).

Although limited, some evidence suggests image-independence for familiar faces in the N250r component. For instance, when the different image of a familiar face is repeated, a N250 response is observed, whereas this effect is smaller than when the same images of the same person are used (Schweinberger et al., 2002). Importantly, after learning, the N250r for previously unfamiliar faces increases and gradually shifts from image-dependent to image-independent repetition effects (Zimmermann & Eimer, 2013). These findings imply that the N250r reflects access to the face representation stored in the FRU. However, it is notable that many of these studies do not directly test the defining feature of familiar face representation—its image independence—and they often do not compare familiar and unfamiliar face recognition directly.

Instead, the N250 familiarity effect reveals a robust and reliable distinction between familiar and unfamiliar faces: familiar faces elicit significantly stronger (i.e., more negative) N250 amplitudes (Tanaka et al., 2006; Andrews et al., 2017; Popova & Wiese, 2023; Gosling & Eimer, 2011; Wiese et al., 2019; Kaufmann et al., 2009). For example, in Gosling and Eimer's (2011) experiment, both famous (i.e., familiar) and unfamiliar faces were presented. The N250 was more negative for famous faces. Interestingly, N250 is also sensitive to face learning. After brief exposure to 20 high-variability images of an unfamiliar face, participants showed a clear N250 effect for that previously unfamiliar face. Notably, even when novel images of the learned face were shown, they elicited a similar N250 response as the repeated images used during learning—suggesting image-independence in the recognition of newly learned faces (Andrews et al., 2017).

However, brief image exposure may not fully reflect real-life face learning. A more ecologically valid study exposed participants to real-life face-to-face interaction and later tested recognition using high-variability images. Despite only 10 minutes of exposure—with little variation in viewpoint or setting—participants were able to recognise the learned faces from highly variable “ambient” images (e.g., different viewpoints, lighting, and expressions) and elicited a clear N250 effect (Popova & Wiese, 2022). These findings further support the idea that the N250 component is a neural marker of familiar face recognition and reflects access to stable face representations stored in the FRU (Wiese et al., 2022).

Following the face recognition stage (N250), later ERP components differentiate between familiarity-based recognition and semantic access to identity-specific information. One such marker is the sustained familiarity effect (SFE), a prolonged posterior negativity brain potential between 400–600 ms at occipito-temporal electrodes after the activation of N250 familiarity effect, which may reflect the continued activation of semantic memory and affective content of the perceived face (e.g., Wiese et al, 2019). In their study, the SFE demonstrated a graded pattern, being largest for personally close and emotionally meaningful faces (e.g., friends), smaller for moderately familiar individuals (e.g., lecturers), and absent for famous celebrities. Since participants were typically familiar with the semantic information of all these groups, this pattern was interpreted as reflecting differences in affective significance, thereby indicating the strongest affective responses to close friends and relatives and no affective responses to celebrities. However, the absence of an SFE for celebrities in Wiese et al. (2019) might partly reflect insufficient familiarity, as they noted that not all participants knew the selected celebrities equally well. Supporting this, Wiese et al. (2022) found that when participants selected their own favourite celebrities, a robust SFE emerged, similar to that for personally familiar faces. In contrast, smaller SFEs were observed for less meaningful but still familiar celebrities, and no SFE was found for unfamiliar ones. These results challenge the earlier conclusion that the SFE is generally absent for celebrities, reopening the question of whether semantic knowledge, affective content, or both are critical to the SFE.

By contrast, the N400 is specifically associated with semantic aspects of person recognition. This centro-parietal negative deflection, occurring approximately 300–600 ms after stimulus onset (Schweinberger, 1996; for review, see Wiese, 2011), is triggered when semantic information about the person is integrated with the visual input. In self-priming studies, when a person’s name precedes their face, priming selectively elicits a more negative N400 rather than an N250r effect (Pickering & Schweinberger, 2003; Wiese et al., 2017). Interestingly, even when the prime is semantically related but not the target’s own name (e.g., the name of the person’s spouse), an N400 effect—but not N250r—is observed (semantic priming). This suggests that the N400 reflects access to the Person Identity Node (PIN) and the broader semantic network of the individual, whereas the SFE reflects sustained familiarity without necessarily involving semantic associations.

Building on the finding that the N250 and SFE components reliably index face familiarity and identity processing, it is also important to consider that these ERP markers could be modulated by ethnicity. A substantial body of research has demonstrated that face recognition is modulated by the ethnicity of both the observer and the face stimulus. This phenomenon, known as the other-race effect (ORE), refers to the finding that individuals recognise faces from their own ethnic group more accurately than faces from other ethnic groups (Young et al., 2012; Wiese et al., 2014). Event-related potential (ERP) studies have been particularly informative in uncovering the neural basis of this effect. Research has shown that ethnicity-related modulations occur at multiple processing stages. During early perceptual processing, other-race faces often elicit a larger N170 component, which is thought to reflect increased difficulty in structural encoding (Wiese et al., 2014; Tüttenberg & Wiese, 2023). At later stages, differences have been observed in components associated with memory encoding and retrieval. For instance, Dm (difference due to subsequent memory) effects, which index successful encoding, are often more pronounced for own-race than for other-race faces starting from approx. 200-300ms, suggesting that more resources are required to encode other-race faces into memory (Lucas et al., 2011; Tüttenberg & Wiese, 2021). Similarly, old/new effects during recognition memory are typically larger for own-race faces, reflecting more detailed retrieval (Herzmann et al., 2011). Critically, these ethnicity-related differences in ERP markers can occur even when behavioural performance is comparable across ethnic groups (Tüttenberg & Wiese, 2023). Given these findings, it is important to consider participant ethnicity as a potential moderating factor in studies of face learning, particularly when the face stimuli are from a single ethnic background.

Multimodal Integration

Notably, so far we have focused on how faces are learned through visual stimuli. However, in real life, we typically do not learn faces in isolation. When we interact with a person, we see their face, hear their voice, and may smell their body odour or the perfume they are wearing. Therefore, we should not limit face learning studies to visual information alone. These multisensory cues may facilitate the formation of a single-person representation (Burton et al., 2011), which can support identity recognition under multisensory conditions.

Visual–Auditory Integration (VAI)

Previous research has focused largely on visual–auditory integration. For example, both auditory and visual information can be used to recognise items such as animals, cars, phones, and human identities (Lewis et al., 2004). Among these, recognition of human social information benefits most from multisensory integration. Auditory input during social interaction typically aligns with facial movement, and both are perceived simultaneously. As a result, facial movements can facilitate speech comprehension (Summerfield et al., 1989), whereas incongruence between visual and auditory cues disrupts perception (McGurk & MacDonald, 1976). Interestingly, although there is a clear double dissociation between face and voice recognition at the neural level—prosopagnosia patients (with deficits in face recognition) show intact auditory recognition, while phonagnosia patients (with deficits in voice recognition) show preserved face recognition (Neuner & Schweinberger, 2000)—these functions appear to interact. For example, in an acquired prosopagnosia patient (SB), impaired face recognition was associated with a compensatory enhancement in voice recognition (Hoover et al., 2010).

The VAI model provides a cognitive framework for the interaction between auditory and visual identity recognition. According to this model (Young et al., 2020), after low-level auditory analysis, structural features specific to the speaker's identity are extracted. These features activate a Voice Recognition Unit (VRU). The more similar the perceived code is to the stored code in the VRU, the more likely it is to be recognised as familiar. VRU activation may then trigger activation of the Face Recognition Unit (FRU) via an identity analysis block, which compares the identity representations already extracted by the VRU and FRU. Rather than integrating raw sensory features, the identity analysis block supports cross-modal matching of person-specific information and may subsequently activate the Person Identity Nodes (PINs), where semantic information about the individual is stored.. However, the existence of this identity analysis block remains under debate; alternative models propose a direct connection between VRU and FRU (Campanella & Belin, 2007). Overall, these models suggest that voice recognition can facilitate face recognition.

Behavioural studies support the integration of VRU and FRU. In voice–face priming studies, participants are briefly presented with the voice of a familiar identity before seeing the corresponding face (or vice versa). These studies consistently find a priming effect: participants respond faster in identity-congruent voice–face and face–voice conditions (Ellis et al., 1997). Notably, this effect only appears when the interval between prime and target is relatively short and disappears in a long interval condition, possibly because in the long interval condition the brain no longer links the auditory and visual inputs as part of the same identity. Furthermore, participants recognise familiar voices more quickly and accurately when they are paired with faces than when heard in isolation (Schweinberger et al., 2007). Similarly, González et al. (2011) found that combining auditory and visual cues enhanced familiar identity recognition compared to voice alone, although no benefit was found compared to face alone. This may be because face recognition is already highly efficient, limiting any additional benefit from voice cues. In contrast, voice recognition is relatively weaker (Stevenage and Neil, 2014) and thus more strongly facilitated by visual information.

However, for unfamiliar identities, the VAI effect is greatly reduced. In studies by Schweinberger et al. (2007) and Ellis et al. (1997), multisensory input sometimes impaired performance—participants took longer to respond in audiovisual (AV) conditions. This discrepancy between familiar and unfamiliar identities may reflect the fact that we encode face and voice together when forming a memory of an identity, since face and voice are typically perceived synchronously. This supports the idea proposed in Section 1.4: visual and auditory information integrate into a single-person representation (Robertson & Schweinberger, 2010; Burton et al., 2011).

ERP studies also provide evidence for audiovisual integration. In Schweinberger, Kloth & Robertson's (2011) experiment, ERPs were analysed under four conditions: (1) familiar face only, (2) familiar voice only, (3) familiar face + familiar voice, and (4) familiar face + unfamiliar voice. They found that the P100 at posterior electrodes (P9/P10, PO9/PO10) was larger in the face–voice conditions (3 and 4) than in the face-only condition. N170 at the same sites (P9/P10) was smallest in the voice-only condition, followed by the face-only condition, and largest in the face–voice conditions. Importantly, differences between the congruent (condition 3) and incongruent (condition 4) pairings emerged from 250–600 ms over central, frontal, and parietal sites (e.g., F4, C3, Cz, P3), with incongruent trials eliciting more negative ERP waves. Critically, since the P100 is associated with lower level perception, which is not face-selective (Rossion and Jacques, 2008) and the N170 is typically associated with early perceptual processing in face recognition (Herrmann et al., 2005; Bentin et al., 1996), the difference between face-only and face–voice conditions likely reflects the basic mechanism of AV processing.

In contrast, differences within the face–voice conditions likely reflect higher-level integration processes. The N250 and N400 components are associated with access to the FRU and PIN, respectively, in face recognition. Thus, ERP differences between congruent and incongruent face–voice pairings in this time window may reflect access to a unified representation that integrates visual, auditory, and semantic information. Similarly, in Joassin et al.'s (2004) study, although audiovisual integration showed behavioural interference (i.e., reduced performance compared to voice-only conditions), the ERP findings still supported AV integration. They observed differences between AV and V+A (asynchronous) presentations at around 110 ms over posterior regions, around 170 ms over posterior regions, and around 270 ms over central regions. They concluded that neural-level integration can occur regardless of whether it facilitates or interferes with behavioural performance.

In summary, findings from behavioural and ERP studies suggest that auditory and visual inputs interact to support the recognition of familiar identities. This integration appears to facilitate recognition when voice and face information are congruent and familiar, but may interfere with recognition under unfamiliar or incongruent conditions. Particularly, in ERP studies, a significant multisensory integration effect occurs from 250ms. These results support the idea that multisensory input contributes to the formation of a unified, person-specific representation, and provide a conceptual foundation for investigating whether similar integrative processes occur between olfactory and visual inputs.

Visual–Olfactory Integration (VOI)

In addition to voices, olfactory cues have also been shown to enrich social information during interaction. Unlike auditory information, which is usually perceived synchronously with facial input, odours can linger in the environment, allowing us to perceive them before, during, or even after a social encounter. These odours can convey information about a person's identity (Lenochova and Havlicek, 2008), age (Yamazaki et al., 2010) or sex (Russell, 1976), and even personality (Sorokowska et al., 2012), enabling olfaction to play a vital role in social communication.

Supporting this, previous studies have demonstrated integration effects between olfactory and facial information in the recognition of affective content. For example, pleasant odours improve the detection of happy faces (Leppänen & Hietanen, 2003); the evaluation of neutral faces is influenced by odour valence, with neutral faces judged as unpleasant when accompanied by unpleasant odours (Cook et al., 2015, 2018); and sex-neutral faces are more likely to be perceived as female under both male and female learning conditions—an effect referred to as the “femininity effect” (Mutic et al., 2016).

Specifically, olfactory cues complement visual information in the recognition of individuals in animals (Jackel & Trillmich, 2003). Olfaction is critical for offspring recognition in sheep. For instance, damaging the main olfactory epithelium significantly impairs a ewe’s ability to recognise her lamb (Lévy et al., 1995). Other forms of olfactory disruption (e.g., bulbectomy or olfactory nerve sectioning) also lead to deficits in offspring identification (Morgan et al., 1975; Bouissou, 1968). Similar findings have been reported in pigs (Jackel & Trillmich, 2003) and goats (Romeyer et al., 1994), suggesting that olfactory information is essential for identity recognition across several non-human species.

Importantly, humans also use olfactory information to identify individuals. Human skin produces a unique body odour that is detectable by others (Nicolaides, 1974). These idiosyncratic and stable signatures—referred to as “odourprints”—can support identity recognition (Havlicek & Roberts, 2009). People can reliably distinguish the body odour of their relatives (e.g., siblings) from that of strangers (Porter et al., 1986) and can even identify their own clothing by smell from among 100 items worn by others (Lord & Kasprzak, 1989). Mothers can recognise their children purely by smell (Russell et al., 1983), and newborns as young as two days old can distinguish their mother’s amniotic fluid from other stimuli (Schaal et al., 1995, 1998).

These “odourprints” may integrate with visual information to facilitate identity recognition. In an odour priming study, participants responded faster to their own face when it was preceded by their own odour compared to other odours (Platek et al., 2004), suggesting integration between olfactory and facial identity information. This may imply that the single-person representation described earlier includes not only the VRU and FRU, but also an “Olfactory Recognition Unit” (ORU), where pre-activation of the ORU facilitates familiar face recognition.

Based on this idea, similar cross-modal facilitation effects should also apply to other familiar faces. However, a follow-up experiment by Platek et al. (2004) failed to replicate the priming effect using the odours of co-workers or relatives. This null result may be due to the small sample sizes used (only 9–12 participants per experiment) and the relatively small overall effect of cross-modal facilitation (Brédart, 2004). Therefore, in the present study, we aim to use a larger sample to more robustly assess the olfactory–visual integration effect.

The neural mechanisms underlying visual–olfactory integration remain largely unexplored. Some insights have been gained from magnetoencephalography (MEG)—a technique that detects small magnetic field changes caused by neural electrical activity (Hansen et al., 2010). In Walla et al.’s (2005) study, participants learned face images under four learning conditions: rose (pleasant), rotten egg (unpleasant), carbon dioxide (pain stimulus), and room air (control). During testing, the same images were presented without odours. Results showed that both pleasant and unpleasant odours improved recognition performance, whereas the pain stimulus decreased it. MEG data showed reduced neural activity at ~260 ms for both pleasant and unpleasant learning conditions.

However, there are several limitations to this study in evaluating cross-modal facilitation. First, participants were tested on the same images used during encoding, which may not reflect true face learning. As discussed earlier, variability—exposure to multiple, differing images—is essential for learning both invariant and variant features and for forming a stable face representation in the FRU. Therefore, this task likely measured picture recognition rather than genuine face learning.

Second, the study did not directly test cross-modal facilitation, as the odour was only present during learning and not during testing. This makes it difficult to determine whether improved recognition performance was due to the general arousal effect of odours or to associative learning between odour and face. In other words, this experiment does not clarify whether participants formed a specific association between the face and the odour, as the odour was not present during the testing phase. The better recognition performance here might be purely due to the attentional arousal elicited by the odours themselves.

To address these gaps, the present study aims to investigate cross-modal facilitation between olfactory and visual inputs and its underlying neural mechanisms using Event-Related Potentials (ERPs). The study includes a learning phase and a testing phase.

To ensure sufficient variability and support stable face representation formation, participants will watch two 12-minute face learning videos that include changes in viewpoint and lighting. Only one identity will be learned for each video. Each video will be accompanied by a different pleasant odour (rose or lilac). In the testing phase, participants will be shown high-variability images of the two learned faces and a control face, each paired with one of the odours used during learning. Participants will be asked to judge the familiarity of each face while their ERPs are recorded.

For the behavioural task, we hypothesise that participants will show significantly higher familiarity ratings when tested with faces previously learned in conjunction with a congruent odour, compared to faces paired with an incongruent odour or novel control faces.

For the ERP predictions, we build upon findings from face recognition and visual–auditory integration (VAI) studies, which highlight the roles of the N250 and SFE components in accessing robust identity representations and associated semantic information. Therefore, we are focusing on established familiarity effects in the N250 and SFE time windows.

We hypothesise that if olfactory cues facilitate face learning, faces learned in the presence of a specific odour and tested with the same (congruent) odour will elicit more negative (i.e., less positive) N250 and SFE amplitudes than those paired with an incongruent odour or novel faces. This would indicate facilitated access to identity-specific neural representations through olfactory–visual integration.

Experiment 1

Method

Participants

A total of 39 participants were recruited from undergraduate and postgraduate students at Durham University to take part in the study. The sample can therefore be characterized as typically adult, falling within the conventional age range for university students (approx. 18-30 years). Based on the 12 experiment versions arising from counterbalancing odour-face pairings across participants, a target of three participants per version was set to ensure an even distribution across the full counterbalancing scheme, resulting in a minimum planned sample size of 36 participants. Five participants were excluded from the final analysis after being identified as outliers in behavioural performance. Outliers were defined as individuals whose mean recognition score for learned faces (congruent and incongruent conditions) was more than 2.5 standard deviations below the sample mean for learned faces, and whose mean recognition score for the novel control face was more than 2.5 standard deviations above the sample mean for novel control faces, resulting in a final sample of 34 participants (20 Asian, 14 Caucasian). Participants were recruited via the university's SONA system and received either SONA credit or an equivalent Amazon voucher (£8 per hour) as compensation. All participants reported normal or corrected-to-normal vision, no known olfactory impairments, no allergies affecting their sense of smell, and no recent illness in the past week that could affect olfactory function. The study received ethical approval from the Durham University Department of Psychology Ethics Committee, and informed consent was obtained from all participants.

Stimuli

Olfactory stimuli

Two natural olfactory stimuli were used in the study: rose essential oil and lilac essence oil. These specific odours were selected based on prior research indicating their pleasant and socially relevant properties (rose: Seubert et al., 2014; lilac: Syrjänen et al., 2018). Their pleasantness was a key factor, as aversive odours could introduce confounding effects related to negative affect and arousal, which are not the focus of the present investigation. Both oils were commercially sourced from Amazon. For each stimulus, 12 drops of the essential oil were added to a cotton wool ball, which was then placed inside a small, sealable round plastic container (diameter: approximately 10 cm; volume: 240 ml) to allow for controlled and consistent presentation across participants.

Face learning video

Participants were required to learn faces through a 12-minute video recorded specifically for the face learning task. Three such face learning videos were created in total, each featuring a different model. All models were young adult females of Caucasian ethnicity, and were fully informed about the purpose of the study. Models received compensation for their participation in the form of an £8 Amazon voucher.

The videos were filmed in a soundproof laboratory against a uniform black background to minimise visual distractions. The model was seated on a chair, with the camera positioned approximately 50 cm in front of them at eye level to ensure consistent framing of the face. Illumination was provided by a photographic white LED light placed approximately 50 cm from the model and elevated to 30 cm above head height.

To elicit natural facial movements and expressions, each model was instructed to respond to a standardised list of open-ended questions (see Appendix 1.1) while maintaining gaze toward a fixed target (crosshair) in front of them. The video was segmented into six two-minute sections to systematically vary visual input. These segments were manipulated by altering the positions of the camera and lighting to introduce within-subject variability (see

Figure 1).

Figure 1

Example of the 6 sections of face learning video.

a. Camera-M, Light-L b. Camera-M,Light-R c. Camera-L, Light- L

d. Camera-L, Light-R e. Camera-R, Light-L f. Camera-R, Light-R

Note. (a-f) Example of the 6 sections of face learning video. (a) Camera is positioned in the middle, the light source is on the left. (b) Camera in the middle, light source on the right. (c) Camera on the left, light source on the left. (d) Camera on the left, light source on the right.

(e) Camera on the right, light source on the left. (f) Camera on the right, light source on the right.

In the first two segments, the camera remained directly in front of the model while the light source was placed at a 45-degree angle to the left (first segment) and to the right (second segment). In the middle two segments, the camera was positioned at a 45-degree angle to the left of the model, while lighting was again alternated between left and right angles. In the final two segments, the camera was positioned at a 45-degree angle to the right of the model, with the light source again alternating between left and right placements. This setup allowed for systematic variation in viewpoint and illumination, which has been shown to enhance face learning (Ritchie & Burton, 2017; Corpuz & Oriet, 2022).

The six segments were edited into a single continuous 12-minute video using Microsoft Clipchamp.

Face images

The face stimuli used in the testing phase were collected from the three models featured in the face learning videos. Each model was asked to provide 50 high-variability ambient photographs depicting their own face. These images were intended to capture natural variations in viewpoint, lighting, facial expression, and appearance (e.g., hairstyle, makeup, clothing), thereby enhancing ecological validity.

All face images were pre-processed to standardise their presentation. Each image was cropped closely around the head to minimise background distractions and then resized to 190 × 285 pixels, corresponding to approximately 5 × 7.5 cm on the screen. To further reduce low-level perceptual differences unrelated to face identity, such as colour (Wiese et al., 2025), all images were converted to greyscale prior to presentation (see Figure 2).

Figure 2

Examples of ambient images from each of 3 identities (IDs).

Note. (a) (b) (c) 10 examples of ambient images from each of 3 identities (IDs).

Experiment design

The face learning experiment involved three facial identities. Each participant was required to learn two of these identities, with the third serving as a novel control condition. Among the two learned identities, one was assigned to the Congruent Identity (ID) condition and the other to the Incongruent ID condition (see Table 1). In the Congruent condition, the identity was paired with the same odour during both the learning and testing phases. In contrast, in the Incongruent condition, the identity was paired with different odours across the two phases. There were three counterbalanced experimental versions in total (see Table 1).

Each learned identity was paired with either a rose or lilac odour. Consequently, the congruent odour used in the experiment was either rose or lilac, resulting in two congruent odour versions (see Table 2). Additionally, since participants were exposed to two learned identities sequentially, the order of learning was counterbalanced such that the congruent identity was learned either first or second. This produced two learning order versions. Altogether, the design included 12 experimental versions: three counterbalanced versions × two learning orders × two odour versions (see Table 2).

Participants were randomly assigned to one of the 12 counterbalanced versions. To populate the design, three participants were initially assigned to each of the 12 versions, resulting in a planned total of 36 participants. However, due to the exclusion of five participants identified as outliers, three additional participants were recruited to maintain statistical power, yielding a final sample of 34 participants. These additional participants were assigned to the 5 versions that had experienced participant loss, namely Versions 1, 5, 7, 8 and 9 to ensure a more balanced final design, though versions 5 and 7 ultimately had two participants each.

Table 1

Versions of experiment setup

	Congruent ID	Incongruent ID	Novel ID
Version 1.	Face (a)	Face (b)	Face (c)
Version 2.	Face (c)	Face (a)	Face (b)
Version 3.	Face (b)	Face (c)	Face (a)

Table 2

All 12 experiment versions

Versions	Experiment version (Face Congruency)	Congruent Odour	Learning Order
1	1 (Face a)	Rose	Congruent 1st
2	1 (Face a)	Lilac	Congruent 1st
3	1 (Face a)	Rose	InCongruent 1st
4	1 (Face a)	Lilac	InCongruent 1st
5	2 (Face c)	Rose	Congruent 1st
6	2 (Face c)	Lilac	Congruent 1st
7	2 (Face c)	Rose	InCongruent 1st
8	2 (Face c)	Lilac	InCongruent 1st
9	3 (Face b)	Rose	Congruent 1st
10	3 (Face b)	Lilac	Congruent 1st
11	3 (Face b)	Rose	InCongruent 1st
12	3 (Face b)	Lilac	InCongruent 1st

Note. Experiment version refers to the assignment of face identities (see Table 1). “Congruent first” means the odour that is congruent with the to- be- tested face was presented during the first learning session; “Incongruent first” means that odour was presented during the second learning session. The incongruent odour was the other scent (rose if lilac was congruent, and vice versa).

Procedure

The study consisted of two main phases: a face learning phase and a subsequent EEG testing phase. The learning phase was conducted in two adjacent laboratory cubicles, each associated with a distinct olfactory condition—either rose or lilac. The odour was administered via a sealed plastic box containing essential oil–infused cotton wool, placed in each room approximately 30 minutes before the participant's arrival to allow the scent to diffuse evenly throughout the space.

Prior to beginning the learning phase, participants completed a brief screening questionnaire to report any current illness, allergies, or other conditions that could impair olfactory function (see Appendix 2). Participants who reported any symptoms likely to affect their sense of smell were excluded from further participation at this stage.

Following the questionnaire, participants entered each of the two cubicles sequentially and rated the perceived intensity of the odour present in each room. They were then assigned to one of the 12 counterbalanced conditions and viewed one of the face learning videos in one of the rooms under the corresponding olfactory condition. Participants were informed that the video would be muted and instructed to focus on memorising the face shown in the video.

After completing the first learning session, participants were given a five-minute break outside the learning cubicles to refresh their sense of smell and minimize the residual effects of the first odour. They then entered the second cubicle to complete the learning session for the other face under the second odour learning condition.

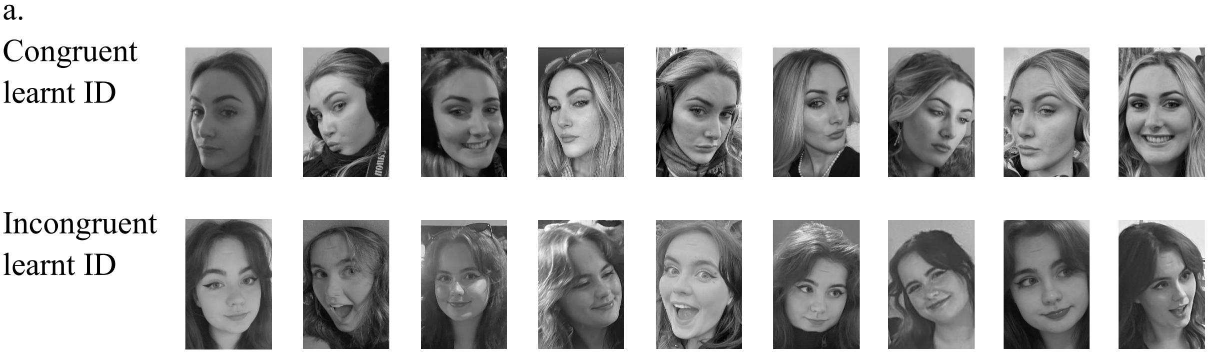
Following the learning phase, participants were taken directly to the EEG laboratory. After approximately 20 minutes of preparation, during which EEG recording equipment was applied, participants were seated in an electrically shielded chamber. Their heads were stabilised using a chinrest positioned 80 cm from a computer monitor. One of the odours previously encountered during the learning phase (either rose or lilac) was presented by placing the corresponding odour box approximately 15 cm below the participant's nose. Participants were again asked to rate the perceived intensity of the odour before the task commenced.

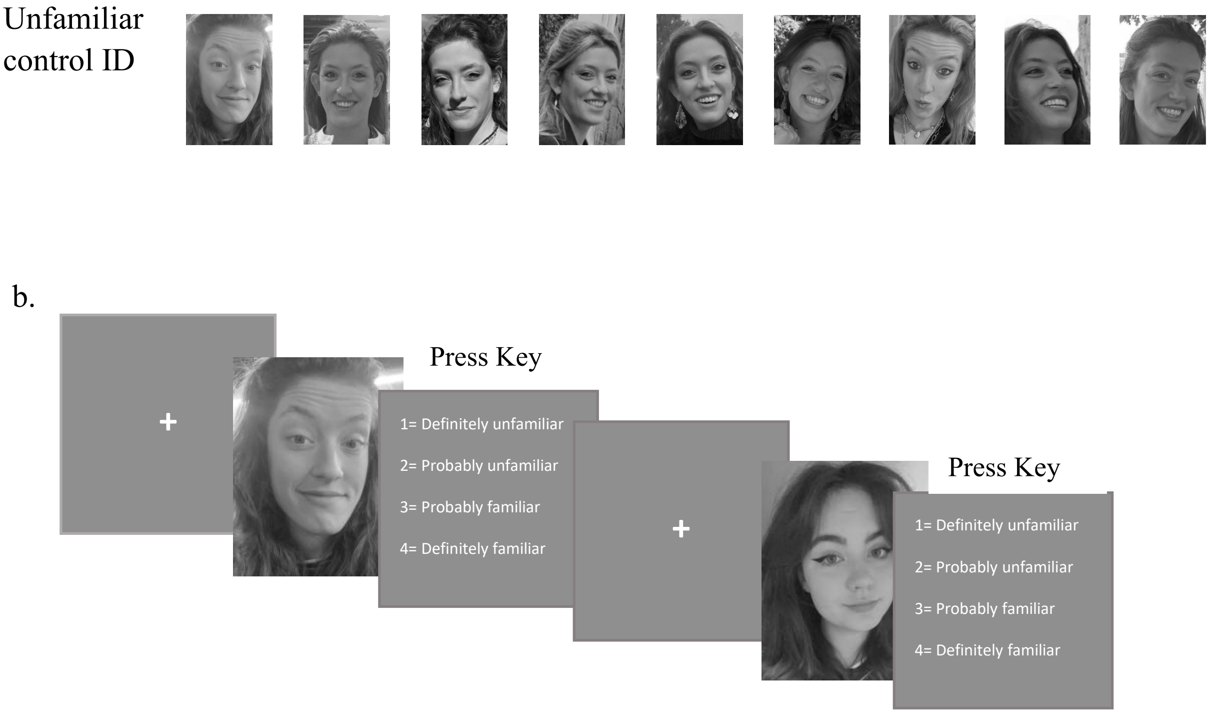
The testing phase consisted of two blocks. In each block, 150 face images were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA), comprising of 50 images from each of the three identities: two previously learned faces (one congruent face and one incongruent face) and one novel face serving as a control. The images were shown in randomised order at a visual angle of 3.6° × 5.4° on a uniform grey background, centrally displayed on the screen for 1000 ms. The inter-stimulus interval varied randomly between 1500 and 2500 ms, with a mean of 2000 ms.

Following each image, participants were instructed to rate the familiarity of the face using a four-point Likert scale presented on the screen. They responded by pressing keys on a keyboard, where ‘1’ indicated “definitely unfamiliar,” ‘2’ indicated “probably unfamiliar,” ‘3’ indicated “probably familiar,” and ‘4’ indicated “definitely familiar.” A 5-minute break was given between the two blocks. The entire EEG testing session lasted approximately 30 minutes (see Figure 3).

Figue 3

Example of 10 images of three IDs and experimental trial structure.





Note. (a) Example of 10 images of three IDs. The first row is the congruent learnt face, the second row is the incongruent learnt face and the third row is the novel face. (b) Example of the experimental trial structure.

EEG recording and analysis

During the experiment, continuous EEG was recorded from 64 scalp sites using sintered Ag/AgCl electrodes (EEGo, ANT Neuro, Enschede, The Netherlands) with a sampling rate of 1024 Hz and a recording bandwidth of DC to 120 Hz. The ground electrode was placed at AFz, and CPz served as the recording reference. Vertical and horizontal eye movements were monitored, and blink artefacts were corrected using Independent Component Analysis (ICA), as implemented in BESA Research 6.3 (Graefelfing, Germany; www.besa.de).

EEG data were band-pass filtered offline using a zero phase shift filter with a 0.1–40 Hz cut-off. Epochs were segmented from −200 ms to 1000 ms relative to stimulus onset and baseline-corrected using the −200 to 0 ms pre-stimulus interval. Artefact rejection was applied using an amplitude threshold of ±100 µV and a gradient criterion of 75 µV/ms. Remaining artefact-free trials were re-referenced to the common average and averaged separately for each experimental condition.

Based on previous research on face learning, ERP analyses focused on the N250 (200–400 ms) and sustained familiarity effect (SFE, 400–600 ms) time windows. Mean amplitudes were calculated over lateral occipito-temporal and temporal sites (FT9/FT10, P9/P10, TP9/TP10). Repeated-measures Analyses of Variance (ANOVAs) were conducted to assess differences between conditions.

Result

Preliminary Analyses

Assumption checks for the ANOVA revealed violations of normality. Shapiro–Wilk tests indicated significant departures from normality for the Congruent condition (p = .014), Incongruent condition (p < .001), and Control condition (p < .001) (see Tables 3, 4 and 5).

Table 3

Descriptive Statistics of the recognition rating with all participants (N = 34).

	Congruent	Incongruent	Control
M	3.50	3.31	1.38
SD	0.41	0.68	0.51

Min.	2.55	1.80	1.00
Max.	3.98	4.00	3.16
P-value of Shapiro-Wilk	0.014	<.001	<.001

Table 4.

Descriptive Statistics of the recognition rating with Asian participants (N = 20).

	Congruent	Incongruent	Control
M	3.53	3.40	1.34
SD	0.43	0.69	0.57
Min.	2.55	2.59	1.00
Max.	3.98	3.97	3.16
P-value of Shapiro-Wilk	0.031	<.001	<.001

Table 5.

Descriptive Statistics of the recognition rating with Caucasian participants (N = 14).

	Congruent	Incongruent	Control
M	3.45	3.17	1.44
SD	0.40	0.66	0.42
Min.	2.59	1.80	1.01
Max.	3.97	4.00	2.34
P-value of Shapiro-Wilk	0.391	0.386	0.036

Given these violations of normality and the relatively small sample size in the current study ($N = 34$), the data were analysed without transformation or correction. Results should therefore be interpreted with caution.

Levene's tests were conducted to assess the assumption of homogeneity of variances. No significant violations were observed: Congruent condition, $F(1, 32) = 0.53$, $p = .47$;

Incongruent condition, $F(1, 32) = 0.01$, $p = .94$; Control condition, $F(1, 32) = 0.26$, $p = .61$. Thus, the assumption of equal variances across groups was met.

However, the assumption of sphericity was not met for the main effect of odour, Mauchly's $W = .78$, $\chi^2(2) = 7.89$, $p = .019$. Therefore, the Greenhouse–Geisser correction was applied in subsequent analyses.

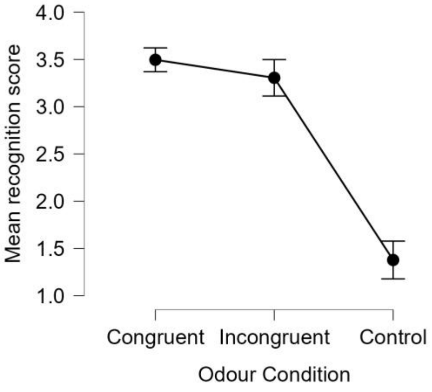
Analysis of Variance

A 2 (ethnicity: between-subjects) × 3 (learning condition: within-subjects) mixed ANOVA was conducted. There was no significant main effect of ethnicity, $F(1, 32) = 0.31$, $p = .58$, $\eta^2p = .01$. However, results of the omnibus test with Greenhouse–Geisser correction revealed a significant main effect of learning condition, $F(1.63, 52.56) = 173.48$, $p < .001$, $\eta^2p = .84$. To test the differences between the three conditions, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at p=0.05/3=0.0167. Planned contrasts indicated no significant difference between the Congruent and Incongruent conditions, M_{diff}= -0.20, 95% CI [-0.44, 0.453], t (64) = 1.64, $p = .105$, *Cohen's d* = -0.37, 95% CI [-0.19, 0.94]. In contrast, participants reported significantly lower scores in the Control condition compared to both the Congruent (M_{diff}= -2.10, 95% CI [-2.35, -1.85], t (64) = -16.89, p <.001, *Cohen's d* = -3.82, 95% CI [-5.12, -2.54]) and Incongruent conditions (M_{diff}= -1.90, 95% CI [-2,15, -1.65], t (64) = -15.25, p <.001, *Cohen's d* = -3.46, 95% CI [-4.65, -2.27]; see Figure 4).

Additionally, the interaction between ethnicity and learning condition was not significant, *F*(1.63, 52.26) = 0.94, *p* = .38, η^2p = .028 (see Figure 5).

Figure 4

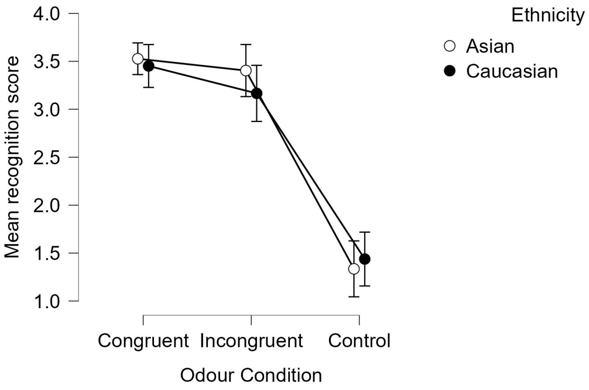
Recognition score across three learning conditions (Congruent, Incongruent and Control)



Note. Error bars represent 95% confidence intervals.

Figure 5

Recognition score of Asian and Caucasian participants across three learning conditions (Congruent, Incongruent and Control).



Note. Error bars represent 95% confidence intervals.

Correlation test

A Pearson correlation was conducted to examine whether the perceived intensity of the Congruent odour was related to the magnitude of the odour congruency effect (i.e., the difference in recognition scores between the Congruent and Incongruent conditions). This analysis was performed to explore if individual differences in odour perception might explain variability in the congruency effect. On average, participants rated the Congruent odour as moderately intense (*M* = 7.37, *SD* = 1.40; see Table 6) on a 10 points scale, ranging from 1 (extremely weak) to 10 (extremely strong). The correlation was not significant, *r*(32) = .15, 95% CI [0.46, -0.20], *p* = .41. This suggests that the difference in recognition scores between the Congruent and Incongruent conditions was not related to the perceived intensity of the Congruent odour.

Table 6

Descriptive Statistics of the perceived intensity of congruent odour.

	ConIntensity	InConIntensity
M	7.37	7.15
SD	1.40	1.78
Min.	4.00	2.00
Max.	10.00	10.00

Event-related Potentials

Figure 6

Grand average Event-related Potential for All Participants at electrodes TP9/TP10 and P9/P10.

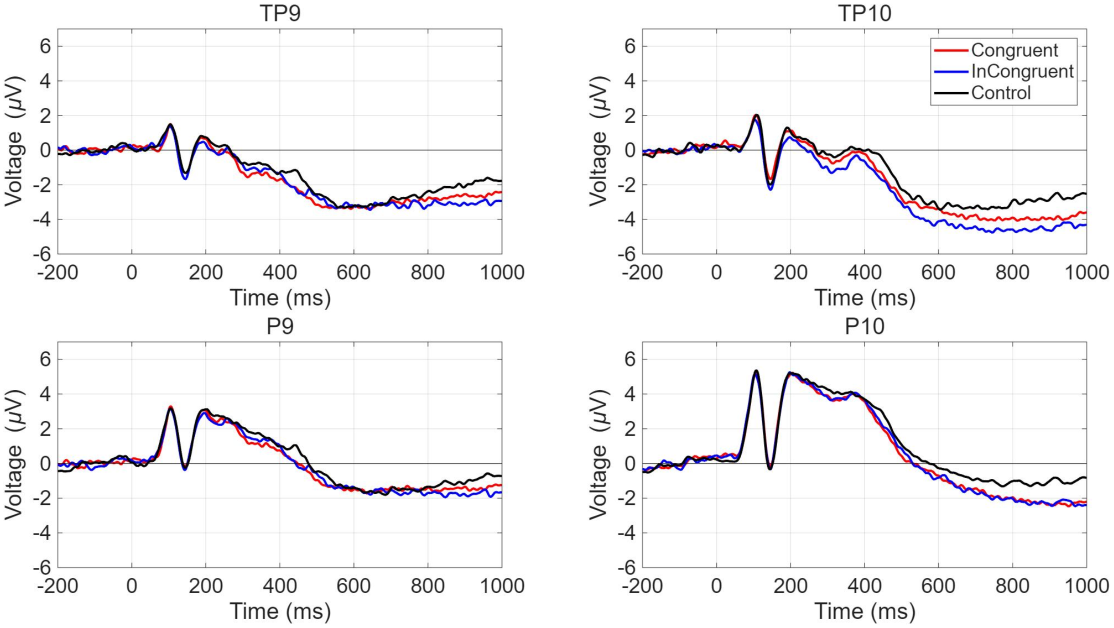


Figure 7.

Grand average Event-related Potential for Asian Participants at electrodes TP9/TP10 and P9/P10.

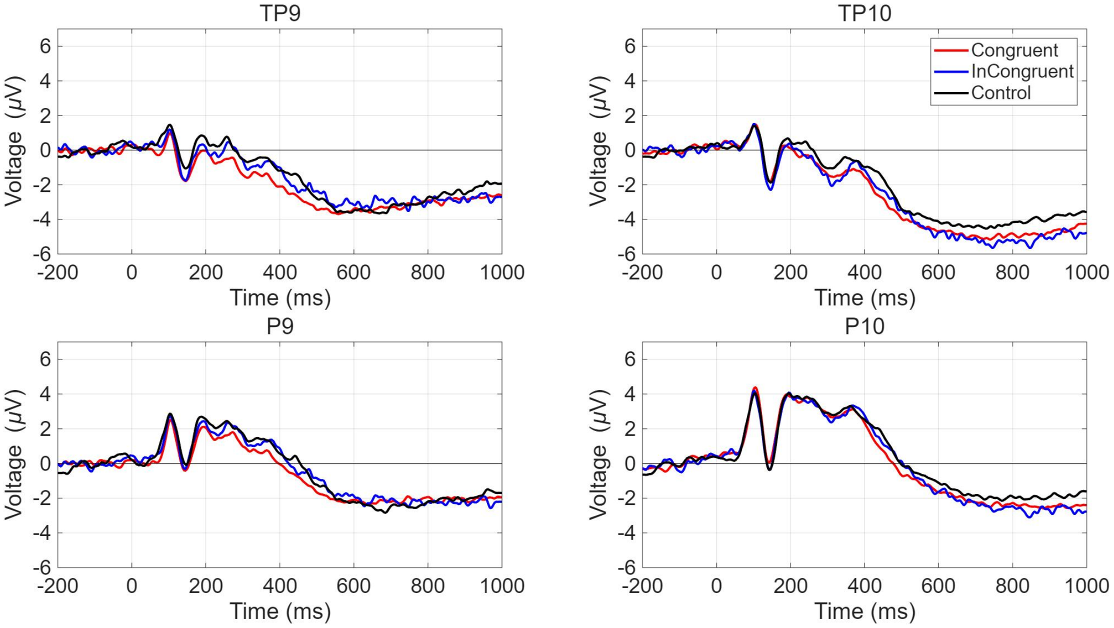


Figure 8

Grand average Event-related Potential for Caucasian Participants at electrodes TP9/TP10 and P9/P10.

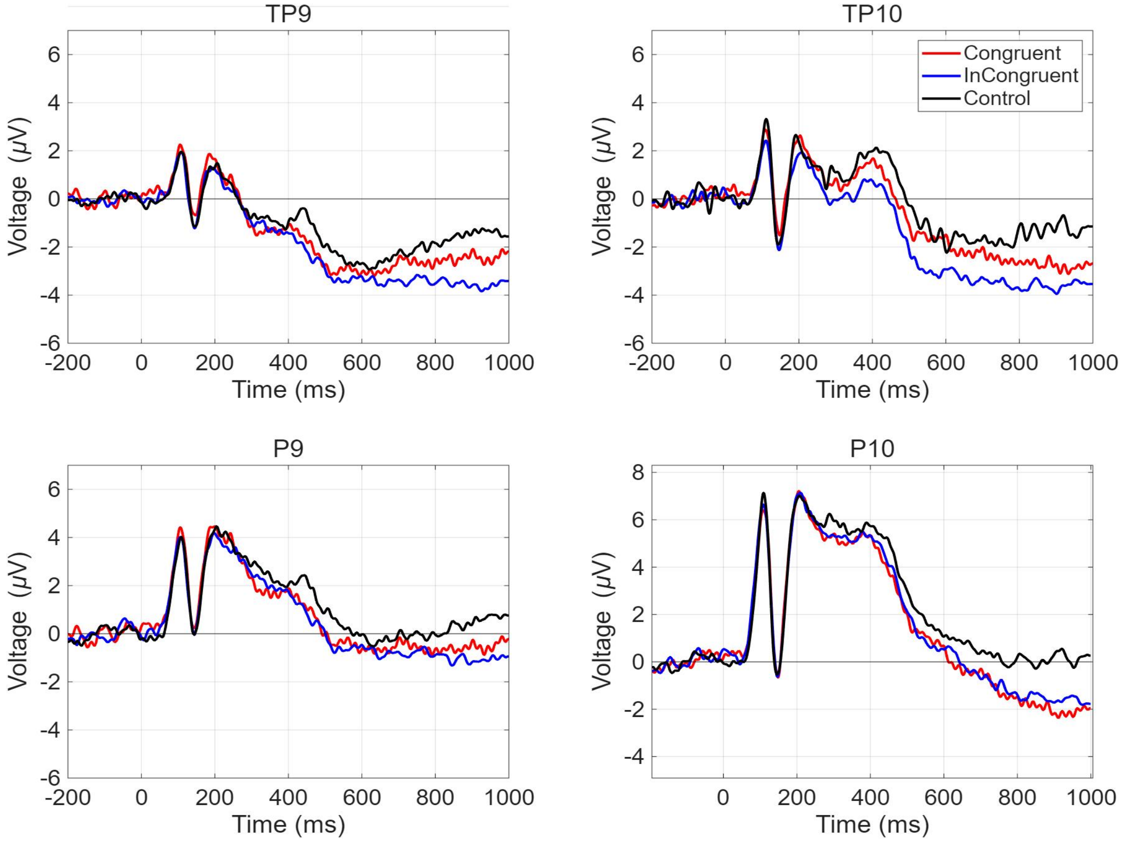


Figure 9.

Difference waveforms (± 95% CI) of congruent and incongruent relative to control conditions for all participants.

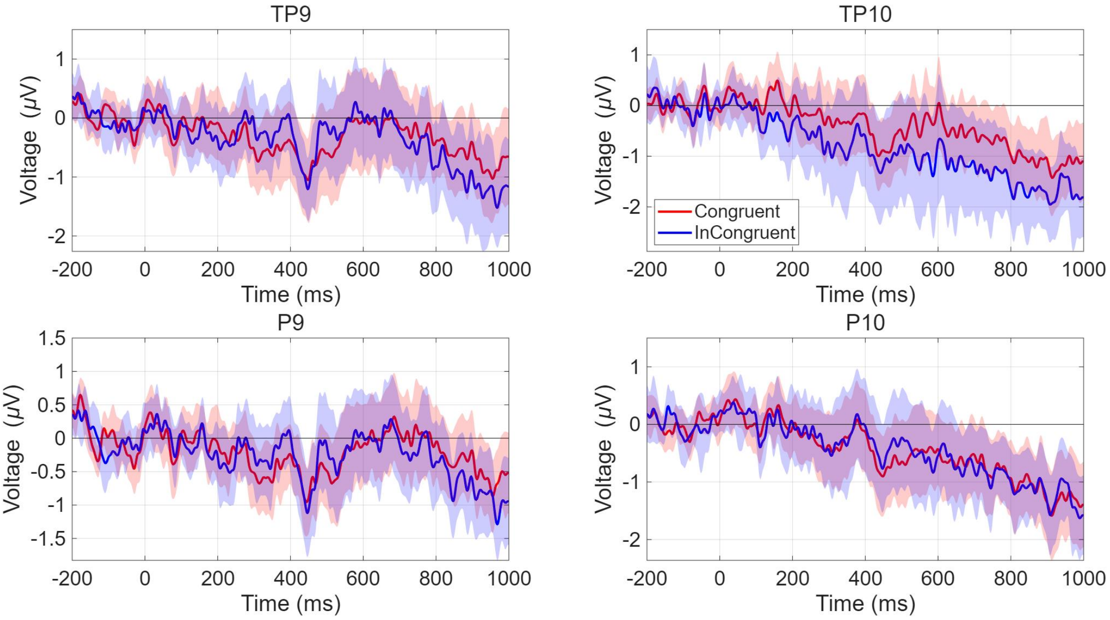


Figure 10

The waveform (± 95% CI) of congruent and incongruent condition for Asian participants

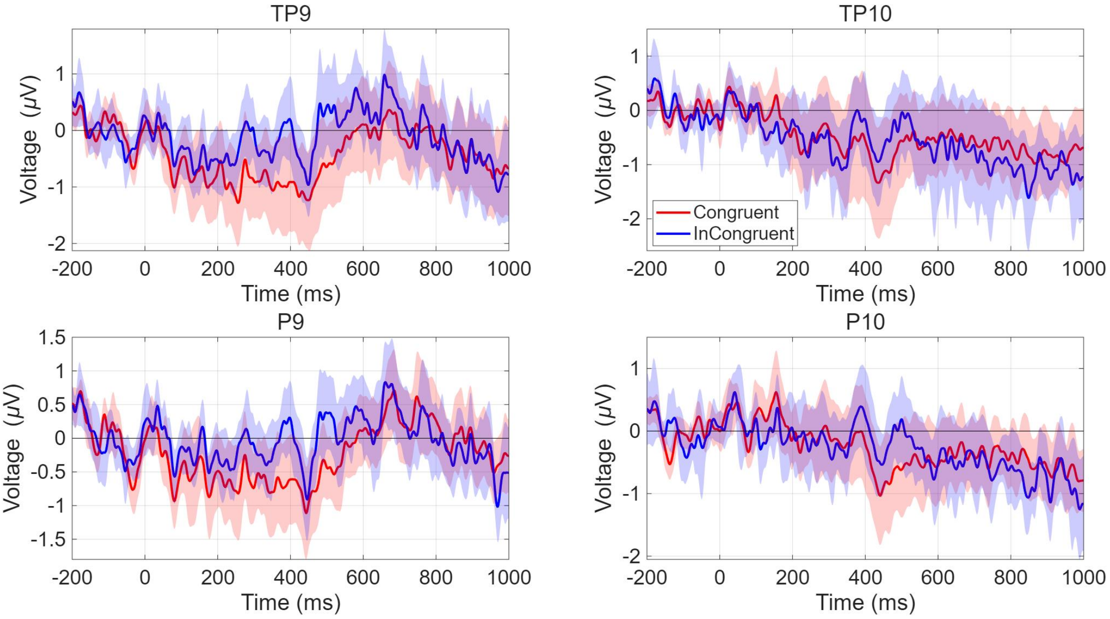
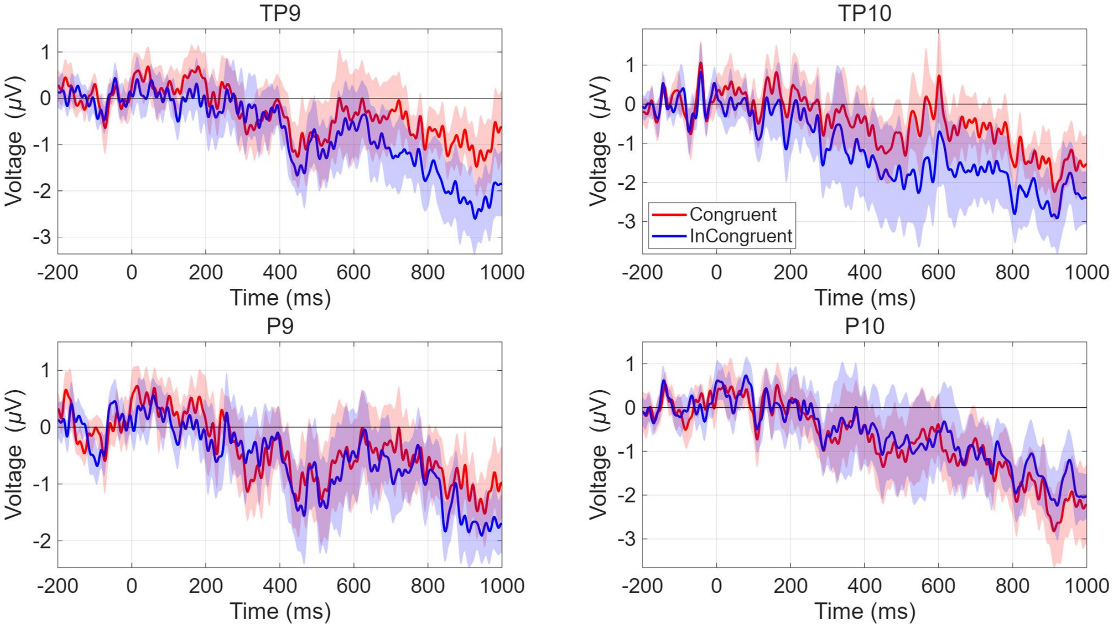


Figure 11

The waveform (± 95% CI) of congruent and incongruent condition for Caucasian participants



The grand average ERP results for all participants are depicted in Figure 6. The grand average ERP results for Asian and Caucasian participants are shown in Figures 7 and 8, respectively. The difference waveforms for congruent and incongruent as compared to control conditions for all participants are shown in Figure 9, and those for Asian and Caucasian participants are shown in Figures 10 and 11, respectively.

The separate 2 (ethnicity: Asian, Caucasian) × 2 (hemisphere: left, right) × 2 (site: P, TP) × 3 (learning condition: Congruent, Incongruent, Control) repeated-measures ANOVAs were conducted for each time window (200–300 ms, 300–400 ms, 400–600 ms, 600–800 ms), with mean ERP voltage as the dependent variable.

Overall, although effects were relatively weak, there was a significant learning effect: the learning conditions (Congruent and Incongruent) produced more negative amplitudes than the Control condition. However, the difference between congruent and incongruent conditions was insignificant across all time windows, except for the Caucasian participants in the 200-300 ms time window. In addition, there was a significant interaction between ethnicity and learning condition, potentially suggesting the presence of an other-race effect.

200–300 ms

At 200–300 ms, a repeated-measures ANOVA on electrodes P9/P10 and TP9/TP10 revealed no significant main effect of learning condition, $F(2, 64) = 2.20$, $p = .12$, $\eta^2 p = .064$.

There was, however, a significant interaction between learning condition and ethnicity, $F(2, 64) = 5.87$, $p = .005$, $\eta^2 p = .16$. Simple main effects analyses showed no significant differences among Asian participants, $F(2, 64) = 2.54$, $p = .39$, but a significant effect among Caucasian participants, $F(2, 64) = 6.19$, $p = .003$. To test the differences between the three conditions for Caucasian participants, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at p=0.05/3=0.0167. Planned contrasts for Caucasian participants revealed that although there was no significant difference between control and congruent condition (M_{diff}= 0.38, 95% CI [-0.29, 1.06], t (26) = 1.17 , p =.253, *Cohen's d* = 0.081, 95% CI [-0.11, 0.27]), the incongruent condition was significantly more negative than the control condition (M_{diff}= 1.08, 95% CI [0.40, 1.75], t (26) = 3.28, p =.003, *Cohen's d* = 0.23, 95% CI [0.003, 0.45]). Additionally, the difference between the Congruent and the Incongruent condition was not significant (M_{diff}= 0.69, 95% CI [-0.017, 1.37], t (26) = 2.11 , p =.045, *Cohen's d* = 0.145, 95% CI [-0.058, 0.35]). Notably, the degrees of freedom for these planned contrasts are derived from the pooled error term of the full factorial model. The calculation is: (Total N - K) × (L - 1), where K is the number of between-subjects groups and L is the number of within-subject conditions. Here, (34 - 2) × (3 - 1) = 64.

300–400 ms

At 300–400 ms, there was no significant main effect of learning condition, $F(2, 64) = 2.50$, $p = .09$, $\eta^2p = .07$. However, there was a significant interaction between learning condition and ethnicity, $F(2, 64) = 4.38$, $p = .016$, $\eta^2p = .12$. Simple main effects analyses revealed no significant differences for Asian participants, $F(2, 64) = 0.68$, $p = .51$. In contrast, Caucasian participants showed significant differences across conditions, $F(2, 64) = 5.37$, $p = .007$, $\eta^2p = .14$. To test the differences between the three conditions for Caucasian participants, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at $p=0.05/3=0.0167$. Planned contrasts showed no significant difference between Congruent and Incongruent conditions, $M_{diff}= 0.39$, 95% CI [-0.47, 1.26], $t\,(26) = 0.94$, $p =.358$, *Cohen's d* = 0.084, 95% CI [-0.162, 0.33]. The incongruent condition, however, was significantly more negative than the control conditions, $M_{diff}= 1.17$, 95% CI [0.31, 2.04], $t\,(26) = 2.79$, $p =.01$, *Cohen's d* = 0.25, 95% CI [-0.026, 0.53]. There was no significant difference between congruent and control condition, $M_{diff}= 0.78$, 95% CI [-0.085, 1.64], $t\,(26) = 1.85$, $p =.075$, *Cohen's d* = 0.17, 95% CI [-0.092, 0.42].

400–600 ms

At 400–600 ms, there was a significant main effect of learning condition, *F*(2, 64) = 4.68, *p* = .013, η^2p = .13. To test the differences between the three conditions, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at *p* =0.05/3=0.0167. Planned contrasts revealed a significant learning effect: the Control condition was significantly more positive than the Congruent (M_{diff}= 0.63, 95% CI [0.13, 1.13], *t* (64) = 2.52 , *p* =.014, *Cohen's d* = 0.180, 95% CI [-0.007, 0.37]) and Incongruent conditions (M_{diff}= 0.69, 95% CI [0.19, 1.19], *t* (64) = 2.76 , *p* =.007, *Cohen's d* = 0.20, 95% CI [0.008, 0.39]). However, Congruent and Incongruent conditions did not differ significantly, M_{diff}= 0.061, 95% CI [-0.44, 0.56], *t* (64) = 0.25 , *p* =.807, *Cohen's d* = 0.018, 95% CI [-0.16, 0.20].

There was also a significant interaction between learning condition and ethnicity, *F*(2, 64) = 6.33, *p* = .003, η^2p = .17. Simple main effects showed no significant differences among Asian participants, *F*(2, 64) = 0.21, *p* = .82, but a significant effect among Caucasian participants, *F*(2, 64) = 6.10, *p* = .007, η^2p = .16. To test the differences between the three conditions for Caucasian participants, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at *p*=0.05/3=0.0167. Planned contrasts revealed that the Control condition was significantly more positive than the congruent (M_{diff}= 1.29, 95% CI [0.32, 2.26], *t* (26) = 2.73, *p* =.011, *Cohen's d* = 0.33, 95% CI [-0.01, 0.67]) and incongruent condition (M_{diff}= 1.53, 95% CI [0.57, 2.50], *t* (26) = 3.25, *p* =.003, *Cohen's d* = 0.39, 95% CI [0.041, 0.74]). Congruent and Incongruent conditions did not differ significantly, M_{diff}= 0.25, 95% CI [-0.72, 1.22], t (26) = 0.52, p =.604, *Cohen's d* = 0.063, 95% CI [-0.26, 0.38].

600–800 ms

At 600–800 ms, there was a significant main effect of learning condition, $F(2, 64)$ = 5.37, p = .007, η^2p = .14. To test the differences between the three learning conditions, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at p =0.05/3=0.0167. Planned contrasts showed that there was no significant difference between Control and Congurent condition (M_{diff}= 0.44, 95% CI [-0.002, 0.87], t (64) = 1.99 , p =.051, *Cohen's d* = 0.14, 95% CI [-0.04, 0.31]). However, Control condition was less negative than the Incongruent conditions (M_{diff}= 0.71, 95% CI [0.27, 1.15], t (64) = 3.25 , p =.002, *Cohen's d* = 0.22, 95% CI [0.038, 0.41]). Congruent and Incongruent conditions did not differ significantly, M_{diff}= 0.28, 95% CI [-0.16, 0.71], t (64) = 1.26 , p =.212, *Cohen's d* = 0.086, 95% CI [-0.087, 0.26].

There was also a significant interaction between learning condition and ethnicity, $F(2, 64)$ = 7.77, p <.001, η^2p = .19. Simple main effects showed no significant differences among Asian participants, $F(2, 64)$ = 0.15, p = .86. However, Caucasian participants showed significant differences across conditions, $F(2, 64)$ = 9.55, p < .001, η^2p = .27. To test the differences between the three conditions, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at p =0.05/3=0.0167.

Planned contrasts revealed no significant difference between Congruent and Incongruent conditions, M_{diff}= 0.59, 95% CI [-0.15, 1.33], t (26) = 1.63 , p =.115, *Cohen's* d = 0.19, 95% CI [-0.13, 0.50]. However, the congruent (M_{diff}= 0.98, 95% CI [0.23, 1.72], t (26) = 2.70 , p =.012, *Cohen's* d = 0.31, 95% CI [-0.035, 0.65]) and incongruent conditions (M_{diff}= 1.57, 95% CI [0.82, 2.31], t (26) = 4.33 , p <.001, *Cohen's* d = 0.49, 95% CI [0.096, 0.89]) were significantly more negative than the control condition.

Discussion 1

The purpose of this experiment was to investigate whether an odour presented during face learning would facilitate subsequent recognition. In summary, the results showed a significant general learning effect in both the behavioural and ERP data, particularly at a later time window (the Sustained Familiarity Effect, SFE; 400-600 ms). However, we did not find evidence of an odour congruency effect—that is, faces were not learned better when paired with a matching odour compared with a non- matching odour. Critically, we observed an other- race effect in the ERP results. Specifically, Caucasian participants exhibited significant N250 (200-400 ms) and SFE (400-600 ms) effects, whereas these effects were absent in Asian participants. Nevertheless, the odour congruency itself was not significant for either group.

Before addressing odour congruency, we first note the clear behavioural learning effect. Memory scores in the learning conditions (Congruent and Incongruent) were significantly higher than in the Control condition, suggesting that even brief exposure through video is sufficient for face learning. Notably, the faces were learned through a video rather than through picture exposures. This finding supports previous work showing that relatively short exposures to within-identity variability can improve recognition performance (e.g., Andrews et al., 2017; Kaufmann et al., 2009; Popova & Wiese, 2023). Furthermore, since highly variable ambient images were used at test - images that are clearly more variable than the learning video - this suggests that participants may have extracted some image-independent features and formed an abstract face representation used for identity recognition.

However, the neural learning effect was only clearly observed at a later time window. In Experiment 1, we focused on the N250 Familiarity effect (200-400 ms) and the Sustained Familiarity Effect (SFE; 400-600 ms). For the N250 Familiarity effect, we examined both an early (200-300 ms) and a late (300-400 ms) time window. Across all participants, the main effect of learning condition was not significant in either window (200-300 ms, p = .12; 300-400 ms, p = .09). This indicates that, overall, there was no reliable neural learning effect between learned and control faces during the N250 time window. Given the clear behavioural learning effect, this finding seems inconsistent with previous N250 studies, which have shown that familiar or learned faces elicit more negative potentials than novel faces within the 200-400 ms window (Tanaka et al., 2006; Andrews et al., 2017; Popova & Wiese, 2023; Gosling & Eimer, 2011; Wiese et al., 2019; Kaufmann et al., 2009).

The absence of an N250 effect after face learning has been reported previously by Tüttenberg & Wiese (2019), who attributed it to a sorting task paradigm that may create only direct image-based links rather than abstract face representations. However, this explanation does not fit our findings for two reasons. First, we observed strong behavioural learning with novel images, confirming successful face learning from the videos. Second, although the N250 effect was not significant overall, we found a significant SFE (400-600 ms), which provides neural evidence that participants successfully learned the faces. Thus, despite the absence of a global N250 effect, other neural markers indicate successful learning.

Two non-mutually exclusive explanations may account for the lack of a significant N250 effect across all participants. One possibility is that the degree of familiarity achieved through a muted video without social interaction was relatively low, leading to smaller N250 effects than those typically observed after real-life encounters (Popova & Wiese, 2023). Another explanation is that each identity was presented 100 times during the test phase, which may have caused even the control faces to become familiar over the course of testing, thereby reducing the difference between conditions (e.g. Itier & Taylor, 2004). Importantly, the N250 was not completely absent in the present study; rather, a significant N250 effect emerged in a subgroup of participants (Caucasians), which will be discussed in the context of ethnicity interactions.

In contrast, for the SFE (400-600 ms), both Congruent and Incongruent conditions were significantly more negative than the Control condition, consistent with prior SFE studies (Wiese et al., 2019; Wiese et al., 2022). Consistent with the proposal by Wiese et al. (2019), this may reflect the integration of affective information with visual representations at this stage.

Having established that learning occurred behaviourally and in the SFE, we now evaluate the central hypothesis of odour congruency. The behavioural test did not demonstrate a facilitation effect of congruent odours. Recognition scores did not differ significantly between Congruent and Incongruent conditions, suggesting that odour congruency did not facilitate face learning. The ERP results mirrored this null finding, showing no significant differences between Congruent and Incongruent conditions across time windows. Together, these results suggest that odour congruency does not provide an additional benefit in face learning. This null finding may be due to several factors. First, compared with auditory-visual integration (Schweinberger et al., 2011; Ellis et al., 1997), olfactory-visual integration may be weaker, leading to smaller effect sizes. Second, in contrast with Platek et al.'s (2004) study, which found congruency effects when participants smelled body odours of familiar individuals, the use of artificial odours here may have reduced ecological validity. Odour-face integration may be more robust in more naturalistic contexts.

Although the congruency hypothesis was not supported, an interaction between ethnicity and learning condition was observed in the ERP data. The other- race effect refers to the phenomenon that people recognise own- race faces more accurately than other- race faces (Young et al., 2012; Wiese et al., 2014; Bar- Haim et al., 2006; Kelly et al., 2007; Hayward et al., 2017). Notably, the current design cannot fully demonstrate the other- race effect because a complete cross- categorisation design requires testing both participant groups with both stimulus types. Behaviourally, the interaction between ethnicity and learning condition was not significant. Therefore, any apparent difference between Asian and Caucasian participants in the congruency effect cannot be interpreted as statistically reliable. The only robust behavioural finding is that both groups learned the faces equally well (no main effect of ethnicity), and odour congruency did not produce a significant behavioural benefit for either group.

The ERP data, however, revealed a clear interaction between ethnicity and learning condition. Although Asian participants showed behavioural learning, they did not exhibit significant ERP learning effects across any time window. In contrast, Caucasian participants showed robust learning effects from 400-600 ms (SFE) for both Congruent and Incongruent conditions, and a significant learning effect for the Incongruent condition from 200-400 ms (N250), which was absent for the Congruent condition. This aligns with previous evidence of the other- race effect in face learning (Hayward et al., 2017). Our findings contrast with Tüttenberg and Wiese (2019), who found ethnicity effects only at N170, not N250. This difference likely results from the limitation of their learning paradigm, where N250 effects appeared only with identical learning- test images, suggesting their task measured image- specific matching rather than genuine face learning that generalises across novel images.

Additionally, as mentioned previously, although the N250 familiarity effect was absent across all participants, Caucasian participants showed a significant N250 familiarity effect for the Incongruent condition, which was absent for the Congruent condition. These findings support previous work on the N250 familiarity effect (Tanaka et al., 2006; Andrews et al., 2017; Popova & Wiese, 2023; Gosling & Eimer, 2011; Wiese et al., 2019; Kaufmann et al., 2009), reflecting the formation of face representations crucial for recognition in Caucasian participants when learning own- race faces. Interestingly, the Congruent condition did not elicit a clear learning effect. Because behavioural and SFE evidence confirmed learning, the absence of an N250 in the Congruent condition cannot be due to failed learning. Instead, it likely reflects a specific modulatory effect of the congruent odour.

One possibility is that odour-face congruency requires more time to integrate, and thus its influence may emerge later in processing rather than during the initial perceptual matching indexed by the early N250. It is also possible that congruent odours created additional associative demands, which delayed rather than facilitated early neural responses. Another explanation is that statistical power was limited in the current study. Although there were 34 participants overall, we analysed the data for Asian and Caucasian participants separately, which resulted in a limited sample size for each group. Future studies with larger samples could help clarify whether congruency influences early stages of face recognition.

The dissociation observed for Asian participants may indicate different encoding strategies for other-race faces. The presence of behavioural learning without corresponding ERP correlates suggests that Asian participants learned Caucasian faces differently from Caucasian participants. Rather than using holistic processing that forms unified representations (Tanaka et al., 2004), they may have relied on more effortful, feature-based strategies. These strategies might not generate strong N250/SFE signals, which reflect automated access to integrated identity representations.

Additionally, it is also possible that the lack of an N250 effect for Asian participants might be due to a ceiling effect of learning, whereby the novel control face became familiar after 100 repetitions during the test phase (Itier & Taylor, 2004). However, this explanation may fail to fully account for the neural results of Asian participants, because, unlike

Caucasian participants, Asian participants also failed to exhibit a Sustained Familiarity Effect (SFE), which reflects access to person-related affective or semantic information. Such information is normally not obtainable purely from image repetition. Since the SFE was clearly shown in Caucasian participants, this highlights the possibility that Asian participants may have used a different strategy during learning and testing.

Nevertheless, due to the limitations of the test phase setup, it remains unclear whether the absence of neural learning effects in Asian participants is attributable to face learning strategy or to experimental limitations. Therefore, to clarify this question, future studies should include novel control face images from multiple different identities rather than from the same person. Notably, one limitation of the present study is that we did not systematically collect precise demographic data, including participants' exact age and gender. Future studies should report these variables to allow better characterisation of the sample and to examine potential effects of age or gender on olfactory–face integration.

In conclusion, although the behavioural congruency effect was not significant in Experiment 1, the ERP results provided two critical insights. First, neural learning effects (N250) were present only for Caucasian participants and only in the Incongruent condition. This suggests that odour congruency may influence face learning in a way that is not simply captured by familiarity ratings, and that ethnicity plays a key role. Second, the complete absence of any N250 and SFE effects in Asian participants indicates that learning other-race faces may rely on different learning strategies.

Given that the behavioural test showed no overall congruency effect, and the ERP interaction pointed to a potential effect only for Caucasians, several methodological factors may have obscured a behavioural effect: (1) the two odours (rose and lilac) were both pleasant floral scents and relatively similar, reducing the distinctiveness of the pairing; (2) the sample of Caucasian participants was small (n = 14), limiting statistical power; and (3) the presence of Asian participants learning other- race faces may have added noise to the overall analysis. Therefore, Experiment 2 was designed to test the behavioural odour congruency effect under more favourable conditions: a larger sample of Caucasian participants only (learning own- race faces), and a more distinctive odour pair (rose vs. sandalwood). If odour congruency genuinely facilitates face learning, we would expect higher recognition scores for Congruent than Incongruent faces in this homogeneous sample, despite the absence of a significant behavioural interaction in Experiment 1. This approach focuses on the population that showed any neural learning effect (Caucasians) and maximises the chance of detecting a behavioural effect by improving odour distinctiveness and sample size.

Experiment 2

Method

Participants

A total of 24 participants were recruited from undergraduate and postgraduate students at Durham University. Specific data on age and gender were not collected due to a methodological oversight; however, the sample can be characterized as typically adult (conventional university student age range, approx. 18-30 years) and mixed-gender. The within-subjects design of the experiment mitigates concerns regarding individual differences in factors such as age and gender. Based on the 12 versions of the experiment arising from counterbalancing odour–face pairings across participants, with each condition presented to two participants, a minimum of 24 participants was required for a fully balanced design. All participants identified themselves as Caucasian (i.e., born and residing in a Caucasian country). Participants were recruited online through Facebook or via leaflets distributed at the university. Each participant received a £10 Amazon voucher as compensation. All participants reported normal or corrected-to-normal vision, no known olfactory impairments, no allergies affecting their sense of smell, and no recent illness in the past week that could have affected olfactory function. Although the final sample is smaller than in Experiment 1, the primary aim was to test the congruency effect in a more homogeneous sample (Caucasian participants only, learning own-race faces) with more distinctive odours, rather than to increase the overall sample size.

Olfactory Stimuli

Two olfactory stimuli were used in the study: rose essential oil and sandalwood essential oil. These odours were chosen with reference to the rose essential oil used in Experiment 1. In that experiment, both odours were floral scents and thus relatively similar to each other. To increase distinctiveness, sandalwood essential oil was introduced. The selection of sandalwood was guided by a small pilot test (n = 3 volunteers) in which several candidate odours (lemon, orange, and sandalwood) were rated for perceived distinctiveness from rose and for pleasantness. Sandalwood was selected because it was rated as clearly distinct from rose (i.e., a non- floral, woody scent) while remaining pleasant, thereby providing a stronger contrast between the two olfactory conditions. Similar to Experiment 1, 12 drops of each essential oil were added to a cotton wool ball, which was then placed inside a small, sealable round plastic container (diameter: 10 cm; volume: 240 ml) to ensure controlled and consistent presentation across participants.

Face Learning Materials, Design, and Procedure

The face learning video, face images, experimental design, and procedure were identical to those used in Experiment 1, except that during the testing phase participants were not required to wear the EEG cap and no EEG recordings were collected. In addition, the new screening questionnaire included a measure of odour pleasantness.

Result

Two of the 24 participants were excluded as outliers based on the interquartile range (IQR) method, resulting in a final sample of 22 participants. Since there should be 2 participants per 12 versions of experiment, the experiment is not fully counterbalanced. There was only 1 participant for version 4 and version 11 (see Table 2)

Assumption checks for the ANOVA revealed violations of normality. Shapiro–Wilk tests indicated significant departures from normality ($p < .05$) for the Congruent ($p < .001$), Incongruent ($p = .003$), and Control conditions ($p = .004$) (see Table 6). Given these violations of normality and the relatively small sample size ($N = 22$), the data were analysed without transformation or correction. Results should therefore be interpreted with caution.

Table 6.

Descriptive Statistics of the recognition rating.

	Congruent	Incongruent	Control
M	3.68	3.48	1.24
SD	0.40	0.48	0.25
Min.	2.32	2.45	1.00
Max.	4.00	4.00	1.73
P-value of Shapiro-Wilk	<.001	.003	.004

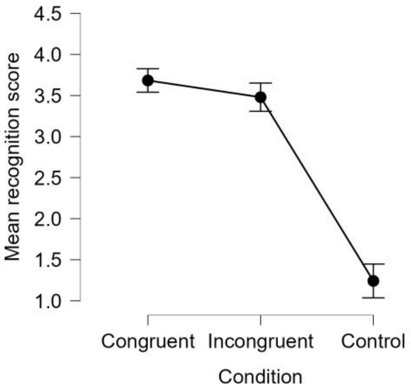
The assumption of sphericity was met, Mauchly's $W = .835$, $\chi^2(2) = 32.75$, $p = .152$; therefore, no correction was required.

A repeated-measures ANOVA was conducted to examine the effects of learning condition (Congruent, Incongruent, Control) on recognition performance. To test the differences between the three conditions, three pairwise planned contrasts were performed. To control for the increased risk of Type I error associated with multiple comparisons, a Bonferroni correction was applied, setting the significance threshold at p =0.05/3=0.0167. There was a significant effect of learning condition, $F(2, 42) = 257.36$, $p < .001$, $\eta^2p = .93$. Planned contrasts indicated that recognition scores in the Congruent (M_{diff}= 2.44, 95% CI [2.20, 2.68], t (42) = 20.45 , p <.001, *Cohen's d* = 6.30, 95% CI [3.79, 8.81]) and Incongruent condition (M_{diff}= 2.24, 95% CI [2.00, 2.48], t (42) = 1.63 , p <.001, *Cohen's d* = 5.78, 95% CI [3.46, 8.10]) were significantly higher than in the Control condition (see Figure 14). A paired-samples t-test further revealed that the Congruent condition ($M = 3.68$, $SD = 0.44$) yielded higher ratings than the Incongruent condition ($M = 3.48$, $SD = 0.48$), $t(21) = 2.16$, p

= .042, *Cohen's d* = 0.46, 95% CI [0.015, 0.90]. This finding should be interpreted with caution, however, as the p-value does not survive correction for multiple comparisons (e.g., a Bonferroni correction for the three comparisons would set the significance threshold at p<.0167). Therefore, the congruency effect observed in Experiment 2 is preliminary and requires replication in a larger sample (see Figure 14).

Figure 14.

Recognition score across three learning conditions (Congruent, Incongruent and Control).



Note. Error bars represent 95% confidence intervals.

Correlation test

A Pearson correlation was conducted to examine whether the perceived intensity and pleasantness of the Congruent odour was related to the magnitude of the odour congruency effect (i.e., the difference in recognition scores between the Congruent and Incongruent conditions). This analysis was performed to explore if individual differences in odour perception might explain variability in the congruency effect. There was no significant difference in odour intensity (M_{diff} = 0.14, 95% CI [-0.58, 0.85], *t* (21) = 0.40 , *p* =.696, *Cohen's d* = 0.085, 95% CI [-0.34, 0.50]) and pleasantness (M_{diff} = 0.18, 95% CI [-0.74, 1.11], *t* (21) = 0.41 , *p* =.687, *Cohen's d* = 0.087, 95% CI [-0.33, 0.51]) between congruent and incongruent condition.

The correlation between the difference in recognition score and the difference in odour intensity was not significant, *r* (21) = 0.13, 95% CI [-0.31, 0.52], *p* = .578. Similarly, the correlation between difference in recognition score and the difference in pleasantness was not significant, *r* (21) = –.10, 95% CI [–0.50, 0.34], *p* = .664.

Table 7.

Descriptive Statistics of the perceived intensity and pleasantness of congruent odour.

	ConIntensity	InconIntensity	ConPleasantness	InconPleasantness
M	7.02	6.89	8.05	7.86
SD	1.33	1.48	1.33	1.49
Min.	4.00	3.00	5.00	5.00
Max.	9.00	9.00	10.00	10.00

Discussion 2

This experiment improved on Experiment 1 by using a larger sample of Caucasian participants (learning own-race faces) and more distinctive odours (rose vs. sandalwood) to investigate olfactory–face integration. The results showed a trend toward a behavioural odour congruency effect: recognition scores were numerically higher for the Congruent condition than for the Incongruent condition, and a paired-samples t-test yielded *p* = .042. However, this result did not survive correction for multiple comparisons (Bonferroni-corrected *p* = .0167). Therefore, the finding does not provide reliable statistical evidence for a congruency effect and should be interpreted with caution. It remains possible that the observed trend reflects a Type I error, and that odour congruency has no genuine effect on face learning even under improved conditions.

Consistent with Experiment 1, recognition scores for both learned conditions (Congruent and Incongruent) were significantly higher than for the Control condition, confirming successful face learning. However, because the difference between Congruent and Incongruent conditions was not statistically significant after correction, we cannot conclude that odour information integrated with visual information to facilitate recognition. The present results do not clearly support the idea of an Olfactory Recognition Unit (ORU); any such interpretation would require a reliable behavioural effect and, ideally, converging neural evidence (e.g., ERP data), which were not collected in this experiment.

An alternative interpretation is that the observed trend, if real, might reflect context-dependent memory (Godden & Baddeley, 1975), where a match between the learning and testing environments enhances retrieval, rather than a person-specific ORU. Because only one identity was learned per odour, the current design cannot distinguish between identity-specific associative learning and general contextual overlap. A follow-up study comparing face–odour with object–odour pairings would help clarify whether any effect is specific to face identity. Moreover, behavioural measures alone cannot determine whether odour congruency influences deeper face-representation processes; future studies should combine behavioural tests with ERP or other neural measures.

The observation that a trend toward congruency occurred only when participants learned own- race faces (Caucasians learning Caucasian faces) raises the possibility that the other- race effect might moderate olfactory–visual integration. However, because Experiment 2 included only Caucasian participants, this remains a speculative hypothesis. A double- dissociation design – with both Asian and Caucasian participants learning both own- and other- race faces – would be necessary to test this possibility directly.

Finally, although Experiments 1 and 2 were designed to be as similar as possible, the learning–testing interval was much shorter in Experiment 2 (5–25 minutes vs. 30–60 minutes). This difference could have enhanced memory performance and may have contributed to the observed trend. Future studies should control the learning–testing interval carefully to rule out this confound. Exposure duration and test format were held constant across experiments.

An additional limitation concerns the measurement of recognition performance. Familiarity ratings were collected on a 4- point scale, and mean scores for learned faces were close to ceiling (Congruent: 3.68; Incongruent: 3.48). This ceiling effect may have reduced sensitivity to detect differences between conditions. Furthermore, response times were not recorded; reaction times can be less susceptible to ceiling effects and may provide a more sensitive measure of recognition difficulty. One additional limitation of the present study is that we did not systematically collect precise demographic data, including participants' exact age and gender. Future studies should report these variables to allow better characterisation of the sample and to examine potential effects of age or gender on olfactory–face integration.

Based on these findings, future research should prioritise the following directions: (a) pre- registered replications with larger samples to determine whether the trend observed here is reliable; (b) the use of natural body odours and face- to- face learning to increase ecological validity; (c) systematic inclusion of both own- race and other- race face stimuli to examine the role of ethnicity; (d) collection of neural measures (e.g., ERPs) to investigate underlying mechanisms; and (e) the use of response time measures or continuous rating scales to avoid ceiling effects.

General Discussion

The present study examined whether odour congruency facilitates face learning, combining behavioural measures across two experiments and ERP measures in Experiment 1. Three main findings emerged from this investigation. First, there was robust evidence of a general learning effect: participants correctly recognised learned faces relative to novel control faces, and this was supported by ERP markers of face learning, particularly at a later time window (the Sustained Familiarity Effect, SFE; 400-600 ms). These findings demonstrate that even brief video-based exposure is sufficient to form face representations that generalise to novel test images.

Second, the hypothesised behavioural odour congruency effect did not appear in Experiment 1 but a trend was observed in Experiment 2, where Caucasian participants showed higher memory scores for faces paired with congruent odours than incongruent odours when learning own-race faces. This difference between experiments might be interpreted as being driven by the other-race effect: participants were more sensitive to olfactory cues when learning own-race faces. Asian participants, by contrast, did not show any congruency effect. Thus, the congruency effect appears to depend on ethnicity and emerges when learning own-race faces.

Finally, an interaction between ethnicity and learning condition was also observed in the neural markers of face learning. Caucasian participants demonstrated clear learning effects at both the behavioural and neural levels (N250 and SFE). In contrast, although Asian participants showed behavioural evidence of learning, they did not exhibit any significant N250 or SFE effects. This dissociation might imply a different face coding strategy for Asian people when learning Caucasian faces. Notably, odour congruency effects were not observed at the neural level for either group.

First, we consider the learning effect. The behavioural data provided clear evidence of successful face learning. Recognition scores for the learned faces (Congruent and Incongruent conditions) were both significantly higher than for Control faces, indicating that 12 minutes of face exposure through video was sufficient to learn a face. This pattern is consistent with prior work showing that even relatively brief exposures can improve recognition performance (e.g., Andrews et al., 2017; Kaufmann et al., 2009; Popova & Wiese, 2023). Importantly, the learning effect observed here is theoretically meaningful.

Based on current theoretical models of face learning and cognition, there exist two distinct yet non-mutually exclusive frameworks: the "Averaging account" and the "Exemplar-based account." According to the Averaging account (Burton et al., 2005), once a face is learned, individual instances are integrated into an abstract representation through a process of averaging. This approach eliminates the variability across images while preserving the critical information for facial recognition which is common to all instances. In contrast, the Exemplar-Based Account (Burton et al., 2016) posits that within-person variability plays a crucial role. Rather than proposing an averaged facial representation, this model suggests that "snapshots" of a face can be reconstructed from parameter values on idiosyncratic dimensions of variability that are learnt for each face. When a new image resembles possible parameter values in this space, the face is perceived as familiar.

The findings of this experiment appear to be better explained by the Averaging Account. There was a difference in variability between the learning and testing phases of the study. During the learning phase, faces were learnt through a 12-minute video characterized by low variability. Although the learning video introduced some variability through changes in viewing angle and lighting, it was recorded in a fixed context with minor to no variation in factors such as hairstyle, makeup or even face expression. In contrast, the test images were highly variable, including differences in lighting, angle, facial expression, background, and clothing. However, participants were able to recognize the identity in high-variability images based solely on learning from the low-variability video. This indicates that they extracted essential facial recognition information from the video and applied it to identify faces in novel contexts. Crucially, since participants could not have obtained "snapshots" of the faces under diverse conditions (e.g. varying environments, expressions, or hairstyles.) from the learning video, it seems that the averaging mechanism may provide a more plausible explanation for the facial learning process observed in the current experiment. Specifically, critical invariant facial features necessary for recognition were abstracted and generalized.

However, the neural learning effect was only clearly observed at a later time window. The present study focused on the N250 (200–400 ms) and the Sustained Familiarity Effect (SFE; 400–600 ms). For the N250, we analysed the early (200–300 ms) and late (300–400 ms) periods. Across all participants, the main effect of learning condition was not significant in either window (200-300 ms; 300-400 ms). Given the clear behavioural evidence of successful learning, this finding appears inconsistent with a substantial body of previous N250 research, which has consistently demonstrated that familiar or recently learned faces elicit significantly more negative amplitudes than unfamiliar faces within the 200–400 ms window (Tanaka et al., 2006; Andrews et al., 2017; Popova & Wiese, 2023; Gosling & Eimer, 2011; Wiese et al., 2019; Kaufmann et al., 2009). The discrepancy between the present null neural effect and the robust behavioural learning effect thus warrants careful consideration.

The absence of the N250 following face learning is not unprecedented (Tüttenberg & Wiese, 2019). They proposed that this absence reflects a limitation of their learning paradigm, which used a sorting task where participants categorized different images by identity. The N250 was absent when novel images were presented at test, suggesting this method may have established only direct links between specific learned and test images rather than forming an abstract face representation.

However, this explanation does not adequately account for the present findings for two key reasons. First, we observed robust behavioural learning effects using novel test images, which confirms that participants successfully learned the faces from the videos despite the absence of a global N250 effect. Second, although the N250 effect was not significant across all participants, we found a significant Sustained Familiarity Effect (SFE; 400–600 ms), which provides clear neural evidence that participants successfully encoded and recognised the learned faces. Taken together, these findings indicate that, even in the absence of a statistically significant N250 effect at the group level, other neural markers – namely the SFE – point to successful face learning.

Two non- mutually exclusive explanations may account for the lack of a significant N250 effect across all participants. One possibility is that the degree of familiarity achieved through a muted video without social interaction was relatively low, leading to smaller and statistically non- significant N250 effects than those typically observed after real- life encounters. For instance, Popova and Wiese (2023) highlighted the role of learning context in face learning. In their study, even a shorter, 10-minute face-to-face interaction produced significant N250 effects. This suggests that direct, socially interactive learning might foster stronger face representations than passive and muted video exposure. Notably, in the current study, participants needed to learn two faces, whereas in Popova and Wiese's (2023) study only one face was learned. Therefore, this additional demand might also reduce the N250 effect.

Another explanation might be attributed to a limitation of the experiment setup. During the test phase each identity – including the novel control face – was presented 100 times. This repeated exposure may have caused the control faces to become familiar over the course of testing, thereby reducing the difference between conditions. Indeed, Itier and Taylor (2004) showed that as few as 10 repetitions during learning phase can elicit N250 familiarity effects, suggesting that 100 presentations may have been sufficient to make the control faces appear familiar, thus obscuring the neural learning effect. Importantly, the N250 was not completely absent in the present study; rather, a significant N250 effect emerged in a subgroup of participants (Caucasians), which will be discussed in the context of ethnicity interactions.

In contrast with the insignificant N250 familiarity effect, we observed a Sustained Familiarity Effect (SFE) for both the Congruent and Incongruent conditions, which reflects the integration of person-related information after the activation of the visual representation and typically occurs between 400–600 ms. In the present study, the learning conditions (Congruent + Incongruent) both elicited more negative amplitudes than the Control condition in this time window, indicating a significant SFE. This finding extends previous work.

So far, most studies investigating the SFE have used highly familiar faces (e.g., close friends, celebrities, lecturers), which provide rich semantic and affective information. In Wiese et al.'s (2019) study, the SFE was suggested to be driven more by affective content than by semantic information. Specifically, the SFE was largest for personally close and emotionally meaningful faces (e.g., friends), smaller for moderately familiar individuals (e.g., lecturers), and absent for famous celebrities. Since participants are typically familiar with the semantic information of all these groups, the graded SFE was interpreted as reflecting differences in affective significance, with the strongest affective responses to close friends and relatives, and no affective responses to celebrities. However, the absence of an SFE for celebrities in Wiese et al. (2019) may partly reflect insufficient familiarity, as they noted that not all participants were equally familiar with the selected celebrities. Supporting this idea, Wiese et al. (2022) found that when participants selected their own favourite celebrities, the

SFE was robust and similar to that elicited by personally familiar faces. Smaller SFEs were observed for less meaningful but still familiar celebrities, while no SFE was found for unfamiliar celebrities. These results challenge the earlier interpretation that the SFE is absent for celebrities and leave open the question of whether semantic knowledge, affective content, or both are critical for the SFE.

In the current face-learning paradigm, identity-specific semantic information was not available, as participants learned novel faces from muted videos. At the same time, affective responses were potentially enhanced by the presence of pleasant odours during learning. Our finding of a significant SFE, with learning conditions eliciting more negative amplitudes than the Control condition, is consistent with Wiese et al.'s (2019) suggestion that affective content plays an important role in the SFE. The present results therefore support the view that the SFE reflects the integration of affective information with visual face representations.

Noteworthy, in contrast with Bojdo et al.'s (2025) study which demonstrate that integrating names with faces consumes a shared pool of cognitive resources, thereby reducing the subsequent SFE. In our study, however, odour-face congruence did not modulate the SFE. This key difference indicates that olfactory-visual integration operates via a distinct neural mechanism, likely one that does not compete for the resources required to sustain later stages of familiar face recognition, unlike semantic integration.

Nevertheless, this study was not specifically designed to test the SFE. In particular, we did not include a control learning condition without odour, which would have allowed us to isolate the role of affective information more directly, or a condition with unpleasant odour to investigate the potential effects of valence. Thus, although our results are consistent with an affective account, further research with better-controlled designs is required.

Additionally, at 600–800 ms, we also found a significant learning effect. However, as this time window was not the main focus of the present study, these results are only briefly noted here.

Taken together, the behavioural and ERP findings provide converging evidence that the present paradigm successfully induced face learning. Participants were able to extract invariant facial features from limited, low-variability video exposure and generalise them to highly variable test images, consistent with averaging accounts of face learning. The ERP results partially corroborated this effect at the neural level, as the neural learning effect was only observed at a later time window (the Sustained Familiarity Effect, SFE; 400-600 ms). The absence of an N250 may be due to the limitation of experiment setup. In addition, the observation of a significant SFE after the N250 suggests that affective information contributes to the integration of person-related knowledge with face representations.

The central aim of the present study was to examine whether odour congruency enhances face learning. Across both behavioural and ERP measures, however, there was no robust evidence that congruent odour–face pairings facilitated recognition compared with incongruent pairings. In the behavioural test of Experiment 1, participants in both learning conditions (Congruent and Incongruent) recognised learned faces better than control faces, but there was no reliable difference between Congruent and Incongruent pairings.

The ERP results mirrored the behavioural patterns. Across time windows, learning effects were observed, but Congruent and Incongruent conditions did not differ consistently. These results suggest that odour–face congruency did not provide a systematic benefit. This finding contrasts with prior research on olfactory–visual integration in person recognition. For instance, Platek et al. (2004) found that participants responded faster when their own face was paired with their own odour compared with other odours.

Several factors may explain these null results. First, olfactory–visual integration may generally be weaker than auditory–visual integration (Schweinberger et al., 2011; Ellis et al., 1997), which is more frequently experienced during social interactions. Voices are consistently paired with faces and facial movements, making them a reliable cue for identity recognition. In contrast, odour cues are rarely consciously perceived during initial encounters with unfamiliar individuals. Reliable odour recognition typically occurs in close relationships, such as with relatives or close friends (Porter et al., 1986; Lord & Kasprzak, 1989; Russell et al., 1983; Platek et al., 2004). Thus, olfactory–visual integration may be relatively small in magnitude, and the modest sample size of the present study (n = 34) may have lacked sufficient power to detect subtle effects. Second, the odours used in Experiment 1 (rose and lilac) were both pleasant but relatively similar, potentially reducing the distinctiveness of odour–face pairings.

Although the congruency hypothesis was not supported in Experiment 1, there was a potential interaction between ethnicity and learning condition. However, this interaction was not statistically significant, so any apparent difference between Asian and Caucasian participants in the congruency effect cannot be reliably interpreted. Because all face models were Caucasian, the non-significant trend toward a larger congruency effect in Caucasian participants may have been due to the limited sample size (n = 14) and the similarity of the floral odours (rose and lilac). To address these limitations, Experiment 2 used a larger sample of Caucasian participants only and more distinctive odours (rose vs. sandalwood).

In Experiment 2, recognition scores were numerically higher for the Congruent condition than for the Incongruent condition ($p = .042$). However, this p-value did not survive correction for multiple comparisons (Bonferroni-corrected $p = .0167$). Therefore, the finding does not provide reliable statistical evidence for a behavioural odour congruency effect. It remains possible that the observed trend reflects a Type I error, and that odour congruency has no genuine effect on face learning even under improved conditions. The results of Experiment 2 should be interpreted with caution and require replication with a larger, pre-registered sample before any firm conclusions can be drawn.

Given the absence of a reliable behavioural effect, the present data do not support the idea of an Olfactory Recognition Unit (ORU) or any specific model of olfactory–visual integration. Any such interpretation would require a statistically robust behavioural effect and, ideally, converging neural evidence (e.g., ERP data), which were not obtained. For completeness, we note two theoretical accounts of how recognition units from different modalities might be connected: a direct connection between modality-specific units (Young et al., 2020; Campanella & Belin, 2007) or convergence at a post-perceptual Person Identity Node (PIN) (Burton et al., 1990). If a reliable congruency effect were to be established in future research, the present ERP finding of a delayed/reduced N250 for the congruent condition in Caucasian participants would be tentatively consistent with the latter (PIN-mediated) account. However, because the behavioural effect was not significant and no neural measures were collected in Experiment 2, this interpretation remains purely speculative.

An alternative interpretation is that any trend toward a congruency effect might simply reflect context-dependent memory (Godden & Baddeley, 1975), where a match between the learning and testing environments enhances retrieval, rather than a person-specific ORU. The current experimental design cannot distinguish between these possibilities, because only one identity was paired with each odour. Consequently, the effect could be attributable to a general contextual overlap rather than to identity-specific olfactory–visual integration. Future studies should compare face–odour pairings with object–odour pairings to determine whether any observed effect is specific to face identity or reflects a broader associative mechanism.

Furthermore, the odours used in the present study were artificial essential oils rather than natural body odours. Although pleasant and controlled, artificial odours lack the idiosyncratic and stable qualities of natural body odours (“odourprints”) that are thought to underlie reliable olfactory identity cues (Havlicek & Roberts, 2009). Future research should therefore use natural body odours to increase ecological validity.

Finally, although the current results hint at a possible other-race effect in olfactory–visual integration (i.e., a trend only for Caucasian participants learning own-race faces), this remains unclear. It could reflect a genuine other-race processing difference or simply greater olfactory sensitivity among Caucasian participants. A double-dissociation design – in which both Asian and Caucasian participants learn both own-race and other-race faces – would be necessary to test this possibility directly.

ERP findings revealed a significant interaction between ethnicity and learning condition. For Caucasian participants, the Incongruent learning condition elicited significant learning effects across all time windows (200–600 ms; N250 and SFE). Although the N250 was absent for the Congruent condition, there was a significant SFE. However, for Asian participants, although they showed behavioural learning, they did not exhibit significant ERP learning effects across time windows. This pattern of results is consistent with previous evidence of the other-race effect in face learning, which has shown that neural responses to own-race and other-race faces differ during recognition (Hayward et al., 2017). However, our findings differ from those reported by Tüttenberg and Wiese (2019), who observed ethnicity-related effects only at the N170 component and not at the N250. This discrepancy likely arises from differences in the learning paradigms used. In their study, significant N250 effects emerged only when the test images were identical to those presented during learning, suggesting that their task primarily measured image-specific matching rather than genuine face learning that generalises across novel views. In contrast, our paradigm used highly variable test images that were different from the learning videos, requiring participants to extract abstract feature information. Therefore, the absence of an N250 effect for other-race faces in our study may reflect a genuine difficulty in forming abstract face representations for other-race faces, rather than a failure to detect image-specific repetition.

Furthermore, as mentioned previously, although the N250 familiarity effect was absent across all participants, Caucasian participants showed a significant N250 familiarity effect for the Incongruent condition, which was absent for the Congruent condition. These findings support previous work on the N250 familiarity effect (Tanaka et al., 2006; Andrews et al., 2017; Popova & Wiese, 2023; Gosling & Eimer, 2011; Wiese et al., 2019; Kaufmann et al., 2009), reflecting the formation of face representations crucial for recognition in Caucasian participants when learning own- race faces.

Interestingly, the Congruent condition did not elicit a clear N250 learning effect for Caucasian participants. Given the robust behavioural learning effect and the significant Sustained Familiarity Effect at the later time window (400–600 ms), we suggest that participants in the Congruent condition successfully learned the faces. Therefore, the lack of an N250 effect specifically in the Congruent condition likely points to an effect of the congruent odour itself.

One explanation is that odour–face congruency may need more time to integrate, so its impact might appear at later processing stages rather than during the initial perceptual matching indexed by early N250. While an incongruent odour might be easily perceived as task-irrelevant noise, a congruent odour would actively engage memory systems in a way that requires additional processing before a stable face representation can be accessed. This extra work takes time which may delay the N250 peak or reduce its amplitude within the typical time window. Supporting this view, an fMRI study on olfactory-visual integration (Gottfried & Dolan, 2003) demonstrated that congruent olfactory cues facilitate visual object recognition. This process is associated with activation in memory-related regions such as the anterior hippocampus which is a structure operating on a longer timescale than direct visual recognition pathways. Consequently, the congruency of odour might increase the associative demand which delay the access of face representation, leading to the delay or reduced N250 effect. Notably, since there was no learning effect for Asian participants under either the congruent or incongruent condition, this effect of the congruent odour might be specific to Caucasian participants only.

Another possibility is that the relatively subtle effect of odour and the use of artificial odour cues, together with the small sample size, limited the statistical power to detect moderately-sized early congruency effects. Because the small sample size for Caucasian participants (N = 14) limited statistical power, future research employing larger samples and high-resolution neuroimaging methods may help determine whether congruency influences the early stages of face processing.

The dissociation observed for Asian participants may reflect the use of different encoding strategies when learning other-race faces. The presence of behavioural learning effects alongside absent ERP correlates suggests that while Asian participants successfully learned and distinguished Caucasian faces, they likely employed different cognitive processes than Caucasian participants. Caucasian participants presumably relied on holistic processing, which integrates facial features into a unified representation (Tanaka et al., 2004). In contrast, Asian participants may have utilized more effortful, strategic, and feature-based processes, such as identifying distinctive characteristics ("this identity has a square jaw"), verbal labeling, or explicit comparison. These analytical approaches might not generate strong signals in the N250/SFE time windows, as these components are specifically linked to automated access to integrated identity representations.

An alternative explanation might be the ceiling effect of learning for Asian participants. Because one important limitation of the experimental design is the lack of a true, unrepeated control identity. During the testing phase, each of the three identities, including the novel control, was presented 100 times. Prior research has shown that far fewer repetitions are sufficient to elicit N250 learning effects. Therefore, it is possible that participants developed a stable N250 response to the "novel" faces over the course of the testing phase itself. Consequently, the absence of learning effect for Asian participants may be partly attributable to this design flaw, rather than the effect of ethnicity.

However, this explanation may fail to fully account for the neural results of Asian participants, because, unlike Caucasian participants, Asian participants also failed to exhibit a Sustained Familiarity Effect (SFE). The SFE is thought to reflect access to person- related affective or semantic information after the activation of the visual representation (Wiese et al., 2019). Such information is normally not obtainable purely from image repetition, because repeated exposure to a face without any personal or semantic context does not generate the kind of rich representation that the SFE is believed to index. Since the SFE was clearly present in Caucasian participants, its absence in Asian participants cannot be easily explained by the repeated presentation of control faces alone. Instead, this dissociation suggests that Asian participants may have used a qualitatively different strategy during learning and testing. Rather than forming integrated, identity- based face representations that support both the

N250 and the SFE, they may have relied on more superficial, feature- based processing that was sufficient for behavioural recognition but did not engage the neural mechanisms indexed by these ERP components. This interpretation is consistent with previous research suggesting that other- race faces are often processed less holistically and with less engagement of identity- specific neural systems (Tanaka et al., 2004; Michel et al., 2006). Thus, the absence of both N250 and SFE effects in Asian participants points to a fundamental difference in how own- race and other- race faces are represented in memory, rather than a simple artefact of repeated stimulus presentation.

Several limitations of the present study should be acknowledged. First, since the effects of odour–visual integration are relatively small, the moderate sample size may have lacked sufficient power to detect subtle effects. Second, the odour stimuli used in the present study were artificial essential oils rather than natural body odours. Although the artificial odours were pleasant and controlled, they lacked the idiosyncratic and stable qualities of body odours that are thought to underlie reliable olfactory identity cues. Third, the face-learning paradigm relied on muted video recordings, which may not evoke the same depth of face encoding as face-to-face social interaction. The relatively small N250 amplitudes observed here may reflect this weaker encoding context and additional learning demand (learning two rather than one new identity). Fourth, a methodological limitation concerns the testing phase: each identity – including the novel control face – was presented 100 times. Prior work shows that far fewer repetitions can elicit N250 familiarity effects (Itier & Taylor, 2004). Therefore, the control face likely became familiar over the course of the experiment, reducing our ability to detect genuine learning effects. A further limitation is that we did not systematically collect demographic data such as the precise age and gender of our participants in Experiment 1 and 2. While our within-subjects design reduces the impact of these variables on our core congruency effect, future studies with larger samples could explicitly test whether factors like gender or age modulate the magnitude of olfactory facilitation in face learning. Finally, the use of only Caucasian face models means the observed other-race effect may not generalise to Asian faces.

Therefore, future studies should address these limitations to more fully investigate the role of odour congruency in face learning. First, increasing the sample size will be important to increase the statistical power to test the relatively small effects of odour–face integration. Second, future studies should use multiple control identities (each presented only a few times). Third, the use of natural body odours and face-to-face learning with interaction, rather than artificial essential oils and learning through muted video, would provide a more ecologically valid test of olfactory identity cues and better capture their role in real-world social interactions. Additionally, using face-to-face learning paradigms may evoke deeper encoding and allow for stronger ERP markers of familiarity, such as the N250 and SFE. Finally, extending the paradigm to include face models from multiple ethnic groups would allow a more systematic investigation of how the other-race effect moderates multisensory integration during face learning.

This study is one of the first to examine odour–face congruency in face learning using both behavioural and neural measures. Although we did not find strong or consistent congruency effects, the results provide useful information about the limits of olfactory–visual integration. First, our findings suggest that olfactory–visual integration may be weaker and slower than auditory–visual integration. This is consistent with the idea that cross-modal integration depends on how often and how reliably different senses are paired in daily life. Second, we observed a dissociation between behaviour and brain responses in Asian participants: they showed behavioural learning of Caucasian faces but no N250 or SFE effects. This suggests that they may have used different encoding strategies for other-race faces, such as feature-based processing, which was enough for behavioural recognition but did not produce the same neural markers as own-race learning. Third, the trend toward a congruency effect in Experiment 2 was only seen in Caucasian participants learning own-race faces. Although this trend was not statistically reliable after correction, it raises the possibility that own-race expertise and social relevance might matter for olfactory–visual integration. This possibility needs to be tested in future studies with larger samples and pre-registered designs.

Overall, these findings suggest that olfactory cues do not enhance face learning. This study has started to examine the conditions under which olfactory information can support face recognition, but future work is necessary to answer the open questions outlined above.

Reference List

Alzueta, E., Melcón, M., Poch, C., & Capilla, A. (2019). Is your own face more than a highly familiar face? *Biological Psychology*, *142*, 100–107. https://doi.org/10.1016/j.biopsycho.2019.01.018

Andrews, S., Burton, A. M., Schweinberger, S. R., & Wiese, H. (2017). Event-related potentials reveal the development of stable face representations from natural variability. *Quarterly Journal of Experimental Psychology*, *70*(8), 1620–1632. https://doi.org/10.1080/17470218.2016.1195851

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and Nurture in Own-Race Face Processing. *Psychological Science*, *17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Begleiter, H., Porjesz, B., & Wang, W. (1995). Event-related brain potentials differentiate priming and recognition to familiar and unfamiliar faces. *Electroencephalography and Clinical Neurophysiology*, *94*(1), 41–49. https://doi.org/10.1016/0013-4694(94)00240-L

Benson, P. J., & Perrett, D. I. (1993). Extracting Prototypical Facial Images from Exemplars. *Perception*, *22*(3), 257–262. https://doi.org/10.1068/p220257

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

Blackwood, D. H. R., & Muir, W. J. (1990). Cognitive Brain Potentials and their Application. *British Journal of Psychiatry*, *157*(S9), 96–101. https://doi.org/10.1192/S0007125000291897

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *10*(5), 527–536. https://doi.org/10.1080/13506280244000168

Brédart, S. (2004). Cross-modal facilitation is not specific to self-face recognition. *Consciousness and Cognition*, *13*(3), 610–612. https://doi.org/10.1016/j.concog.2004.04.001

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*(4), 339–360. https://doi.org/10.1037/1076-898X.5.4.339

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81*(3), 361–380. https://doi.org/10.1111/j.2044-8295.1990.tb02367.x

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, *51*(3), 256–284. https://doi.org/10.1016/j.cogpsych.2005.06.003

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology*, *102*(4), 943–958. https://doi.org/10.1111/j.2044-8295.2011.02039.x

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity From Variation: Representations of Faces Derived From Multiple Instances. *Cognitive Science*, *40*(1), 202–223. https://doi.org/10.1111/cogs.12231

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face Recognition in Poor-Quality Video: Evidence From Security Surveillance. *Psychological Science*, *10*(3), 243–248. https://doi.org/10.1111/1467-9280.00144

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–543. https://doi.org/10.1016/j.tics.2007.10.001

Carbon, C.-C., Schweinberger, S. R., Kaufmann, J. M., & Leder, H. (2005). The Thatcher illusion seen by the brain: An event-related brain potentials study. *Cognitive Brain Research*, *24*(3), 544–555. https://doi.org/10.1016/j.cogbrainres.2005.03.008

Cook, S., Fallon, N., Wright, H., Thomas, A., Giesbrecht, T., Field, M., & Stancak, A. (2015). Pleasant and Unpleasant Odors Influence Hedonic Evaluations of Human Faces: An Event-Related Potential Study. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00661

Cook, S., Kokmotou, K., Soto, V., Wright, H., Fallon, N., Thomas, A., Giesbrecht, T., Field, M., & Stancak, A. (2018). Simultaneous odour-face presentation strengthens hedonic evaluations and event-related potential responses influenced by unpleasant odour. *Neuroscience Letters*, *672*, 22–27. https://doi.org/10.1016/j.neulet.2018.02.032

Corpuz, R. L., & Oriet, C. (2022). Within-person variability contributes to more durable learning of faces. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *76*(4), 270–282. https://doi.org/10.1037/cep0000282

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Eimer, M. (2011). The Face-Sensitivity of the N170 Component. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00119

Ellis, H. D., Jones, D. M., & Mosdell, N. (1997). Intra- and inter-modal repetition priming of familiar faces and voices. *British Journal of Psychology*, *88*(1), 143–156. https://doi.org/10.1111/j.2044-8295.1997.tb02625.x

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of Familiar and Unfamiliar Faces from Internal and External Features: Some Implications for Theories of Face Recognition. *Perception*, *8*(4), 431–439. https://doi.org/10.1068/p080431

Godden, D. R., & Baddeley, A. D. (1975). CONTEXT-DEPENDENT MEMORY IN TWO NATURAL ENVIRONMENTS: ON LAND AND UNDERWATER. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

González, I. Q., León, M. A. B., Belin, P., Martínez-Quintana, Y., García, L. G., & Castillo, M. S. (2011). Person identification through faces and voices: An ERP study. *Brain Research*, *1407*, 13–26. https://doi.org/10.1016/j.brainres.2011.03.029

Gosling, A., & Eimer, M. (2011). An event-related brain potential study of explicit face recognition. *Neuropsychologia*, *49*(9), 2736–2745. https://doi.org/10.1016/j.neuropsychologia.2011.05.025

Gottfried, J. A., & Dolan, R. J. (2003). The Nose Smells What the Eye Sees. *Neuron*, *39*(2), 375–386. https://doi.org/10.1016/S0896-6273(03)00392-1

Hansen, P., Kringelbach, M., & Salmelin, R. (Eds). (2010). *MEG: An Introduction to Methods*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195307238.001.0001

Havlicek, J. (2009). Towards a neuroscience of love: Olfaction, attention and a model of neurohypophysial hormone action. *Frontiers in Evolutionary Neuroscience*, *1*. https://doi.org/10.3389/neuro.18.002.2009

Hayward, W. G., Favelle, S. K., Oxner, M., Chu, M. H., & Lam, S. M. (2017a). The other-race effect in face learning: Using naturalistic images to investigate face ethnicity effects in a learning paradigm. *Quarterly Journal of Experimental Psychology*, *70*(5), 890–896. https://doi.org/10.1080/17470218.2016.1146781

Hayward, W. G., Favelle, S. K., Oxner, M., Chu, M. H., & Lam, S. M. (2017b). The other-race effect in face learning: Using naturalistic images to investigate face ethnicity effects in a learning paradigm. *Quarterly Journal of Experimental Psychology*, *70*(5), 890–896. https://doi.org/10.1080/17470218.2016.1146781

Herrmann, M. J., Ehlis, A.-C., Ellgring, H., & Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission*, *112*(8), 1073–1081. https://doi.org/10.1007/s00702-004-0250-8

Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentzsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology*, *41*(5), 688–701. https://doi.org/10.1111/j.1469-8986.2004.00196.x

Herzmann, G., Willenbockel, V., Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition for own-race and other-race faces.

Neuropsychologia, *49*(11), 3103–3115. https://doi.org/10.1016/j.neuropsychologia.2011.07.019

Hoover, A. E. N., Démonet, J.-F., & Steeves, J. K. E. (2010). Superior voice recognition in a patient with acquired prosopagnosia and object agnosia. *Neuropsychologia*, *48*(13), 3725–3732. https://doi.org/10.1016/j.neuropsychologia.2010.09.008

Itier, R. J., & Taylor, M. J. (2004). Effects of repetition and configural changes on the development of face recognition processes. *Developmental Science*, *7*(4), 469–487. https://doi.org/10.1111/j.1467-7687.2004.00367.x

Jäckel, M., & Trillmich, F. (2003). Olfactory Individual Recognition of Mothers by Young Guinea-Pigs (*Cavia porcellus*). *Ethology*, *109*(3), 197–208. https://doi.org/10.1046/j.1439-0310.2003.00864.x

Jemel, B., Pisani, M., Calabria, M., Crommelinck, M., & Bruyer, R. (2003). Is the N170 for faces cognitively penetrable? Evidence from repetition priming of Mooney faces of familiar and unfamiliar persons. *Cognitive Brain Research*, *17*(2), 431–446. https://doi.org/10.1016/S0926-6410(03)00145-9

Jenkins, R., & Burton, A. M. (2008). 100% Accuracy in Automatic Face Recognition. *Science*, *319*(5862), 435–435. https://doi.org/10.1126/science.1149656

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Joassin, F., Maurage, P., Bruyer, R., Crommelinck, M., & Campanella, S. (2004). When audition alters vision: An event-related potential study of the cross-modal interactions between faces and voices. *Neuroscience Letters*, *369*(2), 132–137. https://doi.org/10.1016/j.neulet.2004.07.067

Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP Correlates of the Acquisition of Face Representations across Different Images. *Journal of Cognitive Neuroscience*, *21*(4), 625–641. https://doi.org/10.1162/jocn.2009.21080

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition*, *172*, 46–58. https://doi.org/10.1016/j.cognition.2017.12.005

Lenochova, P., & Havlicek, J. (2008). Human Body Odour Individuality. In J. L. Hurst, R. J. Beynon, S. C. Roberts, & T. D. Wyatt (Eds), *Chemical Signals in Vertebrates 11* (pp. 189–198). Springer New York. https://doi.org/10.1007/978-0-387-73945-8_18

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and Face Perception: Odors Modulate the Recognition Advantage of Happy Faces. *Emotion*, *3*(4), 315–326. https://doi.org/10.1037/1528-3542.3.4.315

Lévy, F., Kendrick, K. M., Goode, J. A., Guevara-Guzman, R., & Keverne, E. B. (1995). Oxytocin and vasopressin release in the olfactory bulb of parturient ewes: Changes with maternal experience and effects on acetylcholine, γ-aminobutyric acid, glutamate and noradrenaline release. *Brain Research*, *669*(2), 197–206. https://doi.org/10.1016/0006-8993(94)01236-B

Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., & DeYoe, E. A. (2004). Human Brain Regions Involved in Recognizing Environmental Sounds. *Cerebral Cortex*, *14*(9), 1008–1021. https://doi.org/10.1093/cercor/bhh061

Lord, T., & Kasprzak, M. (1989). Identification of Self through Olfaction. *Perceptual and Motor Skills*, *69*(1), 219–224. https://doi.org/10.2466/pms.1989.69.1.219

Lucas, H. D., Chiao, J. Y., & Paller, K. A. (2011). Why Some Faces won't be Remembered: Brain Potentials Illuminate Successful Versus Unsuccessful Encoding for Same-Race and Other-Race Faces. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00020

Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Morgan, P. D., Boundy, C. A. P., Arnold, G. W., & Lindsay, D. R. (1975). The roles played by the senses of the ewe in the location and recognition of lambs. *Applied Animal Ethology*, *1*(2), 139–150. https://doi.org/10.1016/0304-3762(75)90083-8

Mutic, S., Moellers, E. M., Wiesmann, M., & Freiherr, J. (2016). Chemosensory Communication of Gender Information: Masculinity Bias in Body Odor Perception and Femininity Bias Introduced by Chemosignals During Social Perception. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01980

Neumann, M. F., Mohamed, T. N., & Schweinberger, S. R. (2011). Face and object encoding under perceptual load: ERP evidence. *NeuroImage*, *54*(4), 3021–3027. https://doi.org/10.1016/j.neuroimage.2010.10.075

Neumann, M. F., & Schweinberger, S. R. (2008). N250r and N400 ERP correlates of immediate famous face repetition are independent of perceptual load. *Brain Research*, *1239*, 181–190. https://doi.org/10.1016/j.brainres.2008.08.039

Neuner, F., & Schweinberger, S. R. (2000). Neuropsychological Impairments in the Recognition of Faces, Voices, and Personal Names. *Brain and Cognition*, *44*(3), 342–366. https://doi.org/10.1006/brcg.1999.1196

Nicolaides, N. (1974). Skin Lipids: Their Biochemical Uniqueness: Unlike internal organs, the skin biosynthesizes and excretes unusual fat soluble substances. *Science*, *186*(4158), 19–26. https://doi.org/10.1126/science.186.4158.19

Pascalis, O., De Haan, M., & Nelson, C. A. (2002). Is Face Processing Species-Specific During the First Year of Life? *Science*, *296*(5571), 1321–1323. https://doi.org/10.1126/science.1070223

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences*, *102*(14), 5297–5300. https://doi.org/10.1073/pnas.0406627102

Pickering, E. C., & Schweinberger, S. R. (2003). N200, N250r, and N400 Event-Related Brain Potentials Reveal Three Loci of Repetition Priming for Familiar Names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1298–1311. https://doi.org/10.1037/0278-7393.29.6.1298

Platek, S. M., Thomson, J. W., & Gallup, G. G. (2004). Cross-modal self-recognition: The role of visual, auditory, and olfactory primes. *Consciousness and Cognition*, *13*(1), 197–210. https://doi.org/10.1016/j.concog.2003.10.001

Popova, T., & Wiese, H. (2023). How quickly do we learn new faces in everyday life? Neurophysiological evidence for face identity learning after a brief real-life encounter. *Cortex*, *159*, 205–216. https://doi.org/10.1016/j.cortex.2022.12.005

Porter, R. H., Balogh, R. D., Cernoch, J. M., & Franchi, C. (1986). Recognition of kin through characteristic body odors. *Chemical Senses*, *11*(3), 389–395. https://doi.org/10.1093/chemse/11.3.389

Reynolds, J. K., & Pezdek, K. (1992). Face recognition memory: The effects of exposure duration and encoding instruction. *Applied Cognitive Psychology*, *6*(4), 279–292. https://doi.org/10.1002/acp.2350060402

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology*, *70*(5), 897–905. https://doi.org/10.1080/17470218.2015.1136656

Ritchie, K. L., Smith, F. G., Jenkins, R., Bindemann, M., White, D., & Burton, A. M. (2015). Viewers base estimates of face matching accuracy on their own familiarity: Explaining the photo-ID paradox. *Cognition*, *141*, 161–169. https://doi.org/10.1016/j.cognition.2015.05.002

Robertson, D. M. C., & Schweinberger, S. R. (2010). The role of audiovisual asynchrony in person recognition. *Quarterly Journal of Experimental Psychology*, *63*(1), 23–30. https://doi.org/10.1080/17470210903144376

Romeyer, A., Poindron, P., & Orgeur, P. (1994). Olfaction mediates the establishment of selective bonding in goats. *Physiology & Behavior*, *56*(4), 693–700. https://doi.org/10.1016/0031-9384(94)90229-1

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*, *39*(4), 1959–1979. https://doi.org/10.1016/j.neuroimage.2007.10.011

Russell, M. J. (1976). Human olfactory communication. *Nature*, *260*(5551), 520–522. https://doi.org/10.1038/260520a0

Schaal, B., Marlier, L., & Soussignan, R. (1995). Responsiveness to the Odour of Amniotic Fluid in the Human Neonate. *Neonatology*, *67*(6), 397–406. https://doi.org/10.1159/000244192

Schaal, B., Marlier, L., & Soussignan, R. (1998). Olfactory function in the human fetus: Evidence from selective neonatal responsiveness to the odor of amniotic fluid. *Behavioral Neuroscience*, *112*(6), 1438–1449. https://doi.org/10.1037/0735-7044.112.6.1438

Schweinberger, S. R. (1996). How Gorbachev primed Yeltsin: Analyses of associative priming in person recognition by means of reaction times and event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1383–1407. https://doi.org/10.1037/0278-7393.22.6.1383

Schweinberger, S. R., & Burton, A. M. (2003). Covert Recognition and the Neural System for Face Processing. *Cortex*, *39*(1), 9–30. https://doi.org/10.1016/S0010-9452(08)70071-6

Schweinberger, S. R., Kloth, N., & Robertson, D. M. C. (2011). Hearing facial identities: Brain correlates of face–voice integration in person identification. *Cortex*, *47*(9), 1026–1037. https://doi.org/10.1016/j.cortex.2010.11.011

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex*, *80*, 141–153. https://doi.org/10.1016/j.cortex.2015.11.001

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, *14*(3), 398–409. https://doi.org/10.1016/S0926-6410(02)00142-8

Schweinberger, S. R., Robertson, D., & Kaufmann, J. M. (2007). Hearing Facial Identities. *Quarterly Journal of Experimental Psychology*, *60*(10), 1446–1456. https://doi.org/10.1080/17470210601063589

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound Facilitates Visual Learning. *Current Biology*, *16*(14), 1422–1427. https://doi.org/10.1016/j.cub.2006.05.048

Seubert, J., Gregory, K. M., Chamberland, J., Dessirier, J.-M., & Lundström, J. N. (2014). Odor Valence Linearly Modulates Attractiveness, but Not Age Assessment, of Invariant Facial Features in a Memory-Based Rating Task. *PLoS ONE*, *9*(5), e98347. https://doi.org/10.1371/journal.pone.0098347

Sorokowska, A., Sorokowski, P., & Szmajke, A. (2012). Does Personality Smell? Accuracy of Personality Assessments Based on Body Odour. *European Journal of Personality*, *26*(5), 496–503. https://doi.org/10.1002/per.848

Stevenage, S. V., Neil, G. J., & Hamlin, I. (2014). When the face fits: Recognition of celebrities from matching and mismatching faces and voices. *Memory*, *22*(3), 284–294. https://doi.org/10.1080/09658211.2013.781654

Summerfield, R. J., Muehlbauer, F. J., & Short, R. W. (1989). Controlled Environments as an Adjunct to Field Research on Lentils (*Lens culinaris*). V. Cultivar Responses to Above- and Below-average Temperatures during the Reproductive Period. *Experimental Agriculture*, *25*(3), 327–341. https://doi.org/10.1017/S0014479700014848

Syrjänen, E., Wiens, S., Fischer, H., Zakrzewska, M., Wartel, A., Larsson, M., & Olofsson, J. K. (2018). Background Odors Modulate N170 ERP Component and Perception of Emotional Facial Stimuli. *Frontiers in Psychology*, *9*, 1000. https://doi.org/10.3389/fpsyg.2018.01000

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of Preexisting and Acquired Face Representations: The N250 Event-related Potential as an Index of Face Familiarity. *Journal of Cognitive Neuroscience*, *18*(9), 1488–1497. https://doi.org/10.1162/jocn.2006.18.9.1488

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1–B9. https://doi.org/10.1016/j.cognition.2003.09.011

Tüttenberg, S. C., & Wiese, H. (2019). Learning own- and other-race facial identities from natural variability. *Quarterly Journal of Experimental Psychology*, *72*(12), 2788–2800. https://doi.org/10.1177/1747021819859840

Tüttenberg, S. C., & Wiese, H. (2021). Recognising other-race faces is more effortful: The effect of individuation instructions on encoding-related ERP Dm effects. *Biological Psychology*, *158*, 107992. https://doi.org/10.1016/j.biopsycho.2020.107992

Tüttenberg, S. C., & Wiese, H. (2023). EVENT-RELATED brain potential correlates of the OTHER-RACE effect: A review. *British Journal of Psychology*, *114*(S1), 24–44. https://doi.org/10.1111/bjop.12591

Walla, P., Mayer, D., Deecke, L., & Lang, W. (2005). How chemical information processing interferes with face processing: A magnetoencephalographic study. *NeuroImage*, *24*(1), 111–117. https://doi.org/10.1016/j.neuroimage.2004.09.030

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport Officers' Errors in Face Matching. *PLoS ONE*, *9*(8), e103510. https://doi.org/10.1371/journal.pone.0103510

Wiese, H. (2011). The structure of semantic person memory: Evidence from semantic priming in person recognition. *British Journal of Psychology*, *102*(4), 899–914. https://doi.org/10.1111/j.2044-8295.2011.02042.x

Wiese, H., Ablott, E., Bojdo, M. M., & Lidborg, L. H. (2025). Colour is irrelevant for familiar face recognition: Evidence from event-related brain potentials. *Cortex*, *191*, 120–139. https://doi.org/10.1016/j.cortex.2025.07.013

Wiese, H., Anderson, D., Beierholm, U., Tüttenberg, S. C., Young, A. W., & Burton, A. M. (2022). Detecting a viewer's familiarity with a face: Evidence from event-related brain potentials and classifier analyses. *Psychophysiology*, *59*(1), e13950. https://doi.org/10.1111/psyp.13950

Wiese, H., Hobden, G., Siilbek, E., Martignac, V., Flack, T. R., Ritchie, K. L., Young, A. W., & Burton, A. M. (2022). Familiarity is familiarity is familiarity: Event-related brain potentials reveal qualitatively similar representations of personally familiar and famous faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(8), 1144–1164. https://doi.org/10.1037/xlm0001063

Wiese, H., Ingram, B. T., Elley, M. L., Tüttenberg, S. C., Burton, A. M., & Young, A. W. (2019). Later but not early stages of familiar face recognition depend strongly on attentional resources: Evidence from event-related brain potentials. *Cortex*, *120*, 147–158. https://doi.org/10.1016/j.cortex.2019.06.004

Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2014). The Neural Signature of the Own-Race Bias: Evidence from Event-Related Potentials. *Cerebral Cortex*, *24*(3), 826–835. https://doi.org/10.1093/cercor/bhs369

Wiese, H., Komes, J., Tüttenberg, S., Leidinger, J., & Schweinberger, S. R. (2017). Age-related differences in face recognition: Neural correlates of repetition and semantic priming in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1254–1273. https://doi.org/10.1037/xlm0000380

Wiese, H., Schweinberger, S. R., & Kovács, G. (2024). The neural dynamics of familiar face recognition. *Neuroscience & Biobehavioral Reviews*, *167*, 105943. https://doi.org/10.1016/j.neubiorev.2024.105943

Wiese, H., Tüttenberg, S. C., Ingram, B. T., Chan, C. Y. X., Gurbuz, Z., Burton, A. M., & Young, A. W. (2019a). A Robust Neural Index of High Face Familiarity. *Psychological Science*, *30*(2), 261–272. https://doi.org/10.1177/0956797618813572

Wiese, H., Tüttenberg, S. C., Ingram, B. T., Chan, C. Y. X., Gurbuz, Z., Burton, A. M., & Young, A. W. (2019b). A Robust Neural Index of High Face Familiarity. *Psychological Science*, *30*(2), 261–272. https://doi.org/10.1177/0956797618813572

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, *107*(11), 5238–5241. https://doi.org/10.1073/pnas.0913053107

Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater Neural Pattern Similarity Across Repetitions Is Associated with Better Memory. *Science*, *330*(6000), 97–101. https://doi.org/10.1126/science.1193125

Yamazaki, S., Hoshino, K., & Kusuhara, M. (2010). Odor Associated with Aging. *ANTI-AGING MEDICINE*, *7*(6), 60–65. https://doi.org/10.3793/jaam.7.60

Young, A. W., & Burton, A. M. (2017). Recognizing Faces. *Current Directions in Psychological Science*, *26*(3), 212–217. https://doi.org/10.1177/0963721416688114

Young, A. W., Frühholz, S., & Schweinberger, S. R. (2020). Face and Voice Perception: Understanding Commonalities and Differences. *Trends in Cognitive Sciences*, *24*(5), 398–410. https://doi.org/10.1016/j.tics.2020.02.001

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching Familiar and Unfamiliar Faces on Internal and External Features. *Perception*, *14*(6), 737–746. https://doi.org/10.1068/p140737

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and Motivation in Face Recognition: A Critical Review of Theories of the Cross-Race Effect. *Personality and Social Psychology Review*, *16*(2), 116–142. https://doi.org/10.1177/1088868311418987

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., Dong, Q., Kanwisher, N., & Liu, J. (2010). Heritability of the Specific Cognitive Ability of Face Perception. *Current Biology*, *20*(2), 137–142. https://doi.org/10.1016/j.cub.2009.11.067

Zimmermann, F. G. S., & Eimer, M. (2013). Face learning and the emergence of view-independent face recognition: An event-related brain potential study. *Neuropsychologia*, *51*(7), 1320–1329. https://doi.org/10.1016/j.neuropsychologia.2013.03.028

Appendices

Appendix 1. Face learning video questions

Introduction and Background

1. Can you introduce yourself and tell us your current academic program and year?
2. What inspired you to pursue this field of study?
3. Could you share a bit about your background and how you ended up choosing this university/program?

Academic and Research Interests

1. Can you describe your course structure in psychology here at Durham? Are there any unique aspects that you find particularly engaging?
2. Which modules or topics have you enjoyed the most so far, and why?
3. What are your primary research interests?
4. How did you discover your research interest or area of focus?
5. Is there a particular issue or problem in your field that you're especially passionate about?
6. What motivates you to work on this topic?
7. How has studying psychology changed your understanding of human behavior and the mind?

Current or Past Research Projects (Master/PhD student)

1. Can you tell us about any current or past research projects you've worked on?
2. What has been the most rewarding part of your research so far?
3. What challenges have you faced in your research, and how did you overcome them?

Advice and Reflections (Master/PhD student)

1. What advice would you give to someone considering a similar path in research or academia?
2. Looking back, is there anything you wish you had known before starting your academic journey?
3. What's the best piece of advice you've received during your studies?

University Life and Community

1. What is student life like at Durham? Are there any societies, clubs, or activities you're involved in?
2. How would you describe the psychology community here? Do you feel supported by your professors and peers?

3. How has the city of Durham itself influenced your university experience?

Future Plans and Aspirations

1. What are your goals after graduation?
2. Is there a specific career path you're interested in pursuing, or are you still exploring?

Personal Interests and Hobbies

1. What are some of your hobbies or interests outside of your studies?
2. How do you balance your academic work/study with your personal life?
3. Is there a hobby or activity you feel has helped you in your academic journey/study?
4. Do you have a favorite book, movie, or music genre that you'd recommend?
5. What's something surprising or unique about you that most people don't know?
6. If you weren't studying psychology, what would you see yourself doing instead?
7. How do you unwind or recharge after a busy day of studying or work?
8. Are there any sports, arts, or activities that you're passionate about?
9. Do you have a favorite place on campus or in the city where you like to relax or spend time?
10. What's one goal (personal or professional) you hope to achieve outside of your academic work

Appendix 2. Olfaction study questionnaire

Olfaction Study Questionnaire

Purpose: This questionnaire aims to assess individual perceptions of odour intensity and the emotions associated with different scents. Your responses will help us adjust your olfactory experiences.

Participant Information

Do you have any known olfactory impairments (e.g., anosmia, hyposmia)? ☐ Yes ☐ No

Do you have allergies that affect your sense of smell? ☐ Yes ☐ No

Are you currently experiencing a cold, flu, or sinus infection? ☐ Yes ☐ No

Intensity Rating

On a scale of 1 to 10, how strong do you perceive this odour?

(1 = Extremely weak, barely noticeable | 10 = Extremely strong, overwhelming)

- odour Sample 1: ____ /10
- odour Sample 2: ____ /10

How pleasant or unpleasant do you find this odour?
(1 = Extremely unpleasant | 5 = Neutral | 10 = extremely pleasant)

- Odour sample 1: ____ /10
- Odour sample 2: ____ /10

Appendix 3. SONA study summary sheet

SONA Study Summary Sheet

Study name (as it will appear on SONA): Does odour facilitate face learning? An Event-Related Potentials (ERP) study.

Abstract: In current study we investigate whether the combination of olfactory information will induce a stronger N250 effect during face learning.

Detailed description:
To start with, we will test your olfactory function. You will need to smell cotton with a specific essence oil and then rate the intensity of the scent on a scale of 1 to 10.

During the EEG session, you will sit in front of a screen watching a 10-minute face-learning video while wearing an EEG cap. At the same time, we will either provide you with a scent or not (in a control condition) through our olfactory apparatus. Your brain activity will be recorded via electroencephalography (EEG).

When wearing the EEG cap, hair under each electrode is gently separated, and a gel is applied to the scalp to increase the quality of EEG measurement. The gel can be easily washed out after the experiment.

Eligibility Requirements:
You must not have an olfactory disability;
You must not have had any illness that will affect smell function (e.g. Cold, Fluid, Fever, Covid-19, etc.) in past two weeks.
No hairstyles that prevent us from gently separating the hair underneath each electrode (e.g. dreadlocks);
No psychoactive medication;
No skin conditions affecting the scalp (including scalp irritation).

Duration: 1 - 1.5 hour

Credits: 2.0 SONA credits

Researcher(s): Rui Su

Principal Investigator: Prof. Holger Wiese

Appendix 5. Consent form

Consent

Project title: Does odour facilitate face learning? An Event-Related Potentials (ERP) study.
Researcher(s): Rui Su (bcrg35@durham.ac.uk)
Supervisor: Prof. Holger Wiese (holger.wiese@durham.ac.uk)

This form is to confirm that you understand the purposes of the project, what is involved and that you are happy to take part. Please tick each box/select each statement to indicate your agreement.

I confirm that I have read and understand the information sheet and the privacy notice for the above project.	
I have had sufficient time to consider the information, and I am satisfied with the answers I have been given.	
I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
I agree to take part in the above project.	
I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason.	
I understand that anonymised (i.e., not identifiable) versions of my data may be archived and shared with others for legitimate research purposes.	

Signature of Participant .... **Date** ….. / ……. / …..

(NAME IN BLOCK LETTERS) ....

Signature of researcher .... **Date** ….. / ……. / …..

(NAME IN BLOCK LETTERS) ....……….........................

Appendix 6. Participants Information sheet

Participant Information Sheet

Project title: Does odour facilitate face learning? An Event-Related Potentials (ERP) study.
Researcher(s): Rui Su (bcrg35@durham.ac.uk)
Supervisor: Prof. Holger Wiese (holger.wiese@durham.ac.uk)

You are invited to take part in a research study I am conducting as part of my postgraduate dissertation. This study has received ethical approval from the Psychology Ethics Sub-committee of Durham University.
Before you decide whether to agree to take part it is important for you to understand the purpose of the research and what is involved as a participant. Please read the following information carefully and get in contact if there is anything that is not clear or if you would like more information. The rights and responsibilities of anyone taking part in Durham University research are set out in our 'Participants Charter':
https://www.dur.ac.uk/research.innovation/governance/ethics/considerations/people/charter/

What is the purpose of the study?
In this study, we use EEG to investigate whether the combination of olfactory information will induce a stronger brain response to newly learned faces.

Why have I been invited to take part?
You have been invited because you are over 18 and studying Psychology at Durham University.

You should NOT take part if you are left-handed, are taking psychoactive medication, have a brain/neurological condition, have a skin condition or wounds on your scalp, or have any olfactory disability; have had any illness that will affect smell function (e.g. Cold, Flu, Fever, Covid-19, etc.) within past two weeks; take psychoactive medication; have skin conditions affecting the scalp (includes scalp irritation).

Persons not fulfilling these criteria cannot participate for the following reasons: Brain organization and lateralisation of neural processes (such as face recognition) varies with handedness. In order to test an as homogenous group as possible, we will only examine right-handed people. Moreover, both psychoactive medication and brain/neurological conditions affect the EEG signal, which would distort the data collected for the purposes of this study. If people have skin conditions and/or wounds in their scalp, the application of the EEG cap may feel particularly uncomfortable, and they will therefore not be tested.

Do I have to take part?

Your participation is voluntary, and you do not have to agree to take part. If you do agree to take part, you can withdraw at any time, without giving a reason and without any negative consequences such as prejudice or penalty. For more information on your right to withdraw any data identifiable to you in relation to this study, please refer to the privacy notice.

What will happen to me if I take part?

For this study, you will be asked to

- Rate the intensity of scents on scale of 1 to 10 on a questionnaire.
- Attend an EEG+Olfactory session where your brain activity will be monitored when watching a 10-minutes video to familiarize you with a new person while a specific scent is provided.
- The whole experiment will take no more than 90 minutes.

During the experiment, your EEG will be recorded. For that purpose, electrodes mounted in textile caps are placed on the participants' head. The skin under each electrode is rubbed carefully to remove dead skin cells, and the electrode holder is filled with conductive gel. This is necessary to ensure adequate signal quality. The procedure is painless but can feel slightly uncomfortable. The experimenters will reduce any discomfort to you as much as possible. After the experiment, the gel can be easily removed by washing the hair, which is possible in the EEG lab.

Are there any potential risks involved?

There are no particular risks associated with this study

Will there be any physical discomfort?

Preparing participants for EEG is painless but can feel slightly uncomfortable. The experimenters will reduce any discomfort to you as much as possible.

Will there be any psychological discomfort or embarrassment?

There should not be any psychological discomfort or embarrassment.

Can I withdraw from the experiment?

You can request withdrawal of your data, but only during or shortly after the test session. Your data will be anonymised after testing is completed, and it will therefore not be possible to identify them from any of the other data we hold once the experiment is completed. For more information on your right to withdraw any data identifiable to you in relation to this study, please refer to the privacy notice.

Will I receive any compensation for taking part?

You will receive 2.0 SONA Credits

Will my data be kept confidential?

All information obtained during the study will be kept confidential and if the data is published it will

not be identifiable as yours. You will be allocated an anonymous number accessible only by the research team. Your identity will remain anonymous and confidential, will not be disclosed or shared in any circumstances, and will not appear in any publication.
Please read the University's Generic Privacy Notice for important information about how your data will be used.

What will happen to the results of the project?
No personal data will be shared, however anonymised (i.e., not identifiable) data may be used in publications, reports, presentations, web pages and other research outputs. At the end of the project, anonymised data may be archived and shared with others for legitimate research purposes. All research data and records needed to validate the research findings will be stored for 10 years after the end of active data collection.

Who do I contact if I have any questions or concerns about this study?
If you have any further questions or concerns about this study, please speak to the researcher whose contact information is provided at the top of this document. Alternatively, you may contact the Department Ethics Chair.

If you remain unhappy or wish to make a formal complaint, please follow the Department's Complaints Procedure (if you are unable to access this document, please email the Ethics Secretary, who will provide you with a copy).

Thank you for reading this information and considering taking part in this study.

Appendix 7. Participant debrief

Participant Debrief

Project title: Does odour facilitate face learning? An Event-Related Potentials (ERP) study.
Researcher(s): Rui Su (bcrg35@durham.ac.uk)
Supervisor: Prof. Holger Wiese (holger.wiese@durham.ac.uk)

Background and purpose
Face learning is one of the most important social abilities of human beings, enabling us to remember new relationships we form. The N250, a component in event-related brain potentials (ERP) occurring at 250 milliseconds, is recognised as a neural marker of learned faces (Wiese et al., 2019). Notably, scientists typically investigate the face learning process based on visual information alone. However, we live in a multisensory environment with various other types of information around us, such as acoustic, haptic, and olfactory stimuli. Of particular relevance, humans can use olfactory information to identify people (Porter et al., 1986). Despite these

findings, little is known about a potential multisensory effect of olfactory stimuli on face learning. In this project, I am hoping to investigate whether the combination of olfactory information will induce a stronger N250 effect during face learning.Your data will help me to study this.

Withdrawing your data

You can request withdrawal of your data, but only during or shortly after the test session. Your data will be anonymised after testing is completed, and it will therefore not be possible to identify them from any of the other data we hold once the experiment is completed.

How your data will be used

In writing up the study, all data will be anonymised and your personally identifiable data will not be available to anyone outside the research team. Your anonymised data may be archived for verification of the research results.

Questions about the ethics process

If you have any further questions or concerns about this study, please speak to the researcher whose contact information is provided at the top of this document. Alternatively, you may contact the Department Ethics Chair.

If you remain unhappy or wish to make a formal complaint, please follow the Department's Complaints Procedure (if you are unable to access this document, please email the Ethics Secretary, who will provide you with a copy).

Further support

If you have experienced any psychological distress as a result of this study, please reach out to the University Counselling and Mental Health Service (for Durham University students only): Phone: 0191 334 2200; email: cmh.service@durham.ac.uk