MASTERS BY RESEARCH

# Visualising the self-face representation via behavioural image reconstruction

Vinsome, Michael

*Award date:*
2024

*Awarding institution:*
Bangor University

Link to publication

# Visualising the self-face representation via behavioural image reconstruction.

Michael Vinsome

Department of Psychology, Bangor University

Dr. Maister, Prof. Downing

30/09/2023

‘I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

I confirm that I am submitting this work with the agreement of my Supervisor(s).’

*M.Vinsome*

‘Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw’r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o’r blaen ar gyfer unrhyw radd, ac nid yw’n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.

Rwy'n cadarnhau fy mod yn cyflwyno'r gwaith hwn gyda chytundeb fy Ngoruchwyliwr (Goruchwylwyr)

## Abstract:

This project aimed to produce a method of visualizing long-term perceptual memories of faces, to visualize the self-face representation. It aimed to produce the method, test it via the use of an ‘ideal theoretical observer’ using computer vision as initial proof of concept and then pilot with human subjects to assess feasibility. Results were initially promising with images produced being significantly more similar to targets than non-targets. Unfortunately, this finding was not mimicked in participant data leading to alternative approaches to data collection being trialled as well as test-retest validity measures to be run. Both initial test-retest of the original method as well as the alternative approach devised produced only poor non-significant correlations between participants. It remains an open question as to whether the images produced contain a valid ‘self’ like quality to participants.

**Visualising the self-face representation via behavioural image reconstruction.**

The following work aims to produce a method of visualizing long-term perceptual memories of faces, to visualize the self-face representation.

The following work aims to produce a novel application of a method designed to visualise the self-face representation, which is thought to reside in long-term visual memory (Maister et al., 2021). Most works relating to the self-face focus upon self-face perception. Currently only two published works take a data-driven approach to reconstructing the self-face representation (Maister et al., 2021; Moon et al., 2020), but unfortunately the relatively poor quality, noisy visual representations limit their use. The method produced reported here is data-driven and aims to produce images of a higher quality than those produced by the aforementioned methods. This will be predominantly assessed by determining if the method produces images which are more similar to the participants own face than would be expected by chance.

The concept of the self refers to an individual's sense of identity, consciousness, and personal uniqueness. It encompasses one's thoughts, emotions, perceptions, and experiences, forming a cohesive and subjective understanding of who they are as a person. This work will focus on one very particular aspect of the self, namely, the self-face.

In the most literal sense, the ‘self-face’ is the face which belongs to the individual upon whose head and shoulders it is mounted. It is theorized here, however, that the representation of this face is often subject to myriad factors that may influence the internal observer’s objectivity. The ‘internal observer’ refers to the owner of the self-face, and ‘representation’ refers to the concept that most closely corresponds to what the individual perceives when viewing an image of oneself through the mind's eye.

It is axiomatic that the individual's self, of which the self-face representation must surely be a component, is responsible for much of an individual's perspective. As such, the self-face representation is likely to play a pivotal role in how the individual interacts and responds to the world surrounding them at a given moment. Equally, the qualities associated with the self-face representation are also likely to impact the individual's inner world in a significant manner, influencing their thoughts and desires in an abstract fashion.

The self-face representation, as a component of a broader multidimensional 'self,' is likely to play an important role in perspective, thoughts, and desires. The image of one's own face is a particularly distinctive characteristic of self. It is well documented that the mere act of being able to recognize oneself by way of reflection is a critical component of self-awareness, as best exemplified by the rouge test (Gallup, 1970; Calmette & Meunier, 2023), which posits that infants only begin to develop self-awareness and, by proxy, a sense of selfhood after being able to accurately recognize themselves in the mirror.

Although numerous claims have been made of other species also being able to perform the act of mirror self-recognition, such as that garnered by the rouge test, notably evidence has only been consistently reproduced in humans and higher apes (Gallup & Anderson, 2020). Interestingly, however, it appears the ability seems to hold some relationship with the sociability of the animal in question, with more social animals appearing, on average, to present as more 'self-aware' than more solitary creatures (Lei, 2023).

Further research also suggests that the ‘social self’ is recruited in the act of self-face recognition (Sugiura et al., 2012). In this context, the social self can be simply defined as the representation of self, which is seen through the eyes of others; in other words, a part of an individual's self-concept that is derived from their interactions and relationships with others. It encompasses how one perceives themselves in a social context, including how they believe others see and evaluate them. Research suggests that the social self can influence and facilitate self-face recognition in several ways. Perhaps one of the most prominent ways, however, is by aiding the recruitment of ‘self-referential processing’ (Knyazev, 2013). The social self itself requires a sense of self-awareness and self-reference, which has been theorized to lead to enhanced processing of self-related information such as facial features. As such, when viewing an image of one’s own face, the brain is likely primed to engage in self-referential processing, which in turn will facilitate recognition (Sui & Humphreys, 2015).

Building upon this, growing evidence suggests that the self-face representation in its entirety is a somewhat malleable construct, with some lines of research suggesting it is possible to temporarily embody the face of another. Evidence for this stems from the ‘Enfacement illusion’, (Tajadura-Jiménez et al., 2012) in which participants are shown footage of someone else’s face being touched synchronously with their own. This appears to temporarily adjust the self-face representation so that they accept images, which are morphed to include a slightly higher proportion of the other's face (as produced by image morphing software) as their own. Due to this finding, it has been theorized that the representation of the self-face effectively ‘updates’ each time the participant looks into the mirror and experiences the accompanying synchronous sensorimotor stimulation, which the illusion is based upon.

However, it is not just the self-face, which can be manipulated via visuotactile illusions such as the Enfacement illusion. Changes to bodily self-consciousness have also been reported as a result of both the rubber hand (Botvinick & Cohen, 1998) or full-body illusions (Petkova et al., 2011), with individuals reporting a change of either full body ownership or hand ownership when exposed to such paradigms, perhaps further demonstrating the malleability of bodily self-consciousness (Golaszewski et al., 2021). Notably, however, findings pertaining to the rubber hand illusion are disputed, with some sources claiming it is largely the result of confounding variables. One paper claims that, whilst controlling for ‘imaginative suggestion,’ results demonstrating experiences of ‘ownership’ are likely to occur just 4% of the time (Roseboom & Lush, 2022).

The above research demonstrates that the self-face does not appear to behave as a fixed object but dynamically interacts with the world around it. Studies have further demonstrated that we attribute overarching abstract concepts, such as personality, to our perception and representation of other individuals’ faces. (Oh et al., 2021) have successfully shown that when participants believe individuals' personalities to be similar, they also rate their faces as much more similar in appearance. This work strongly suggests that much of face perception may be biased by knowledge of the individual, which is not related to the immediate physical characteristics of the face being perceived.

Additional evidence of the influence of person knowledge on perception can be garnered from the derogation effect (Meyer et al., 2011). In this context, the derogation effect refers to the finding that individuals in committed heterosexual relationships demonstrate a tendency to give lower scores of attractiveness to faces, which have been empirically proven to score highly in attractiveness, than those participants who are not currently involved romantically (Karremans et al., 2011). Overall, both these findings suggest that much of the face perception of others is not an objective act but heavily biased by additional non-perceptual factors.

However, up until very recently, most research into the self-face has focused predominantly on the direct perception and processing of the self-face. Consequently, significantly less is known about how the self-face is represented in memory and what effects beliefs and attitudes might have on these representations. In part, this is due to certain key methodological issues. Several recent lines of inquiry, however, do claim to provide som insight into how the self-face is represented in memory in human subjects.

One such method is that of ‘reverse correlation,’ which is a data-driven method that attempts to reconstruct mental representations of images (Dotsch & Todorov, 2012). In brief, the reverse correlation protocol, when adapted for faces, works by adding patterns of random noise onto a base face image, which is usually the average face of a given stimuli set. This results in a slightly different image each time. Following this, the inverse of the same noise is superimposed onto the same base face image (in essence, each black pixel is now white and vice versa) which results in a different image again. Participants are then asked to choose which of the pair of images best represents the attribute of interest, in the case of this thesis, the self-face. This process is generally performed many times across multiple trials, in a structured fashion with many pairs presented to each participant. From this data a classification image is computed for each individual, this is achieved by creating an average of all selected images' noise and then overlaying that pattern onto the base image. Before the noise is superimposed however, it is scaled to fit the base image so that the smallest and largest pixel intensities are matched to the base image pixels.

Finally, given that representations of others' faces are influenced by our beliefs and attitudes towards them (Oh et al., 2021; Oliveira et al., 2019), it is plausible to suggest that self-face representation is also likely influenced far beyond the purely physical dimension. Adding to this, work by (Maister et al., 2021) has suggested that, using the reverse correlation method outlined above, individuals exhibiting higher social self-esteem produce more accurate 'self-portraits,' the term used for the reverse correlation images thought to reflect the participants' self-face representations. Furthermore, it appeared that individuals' own beliefs about their own personality traits (as measured by the Big 5 personality questionnaire, which has been demonstrated to assess the prevalence of a variety of traits individuals exhibit (Wirawani Kamarulzaman & Prof Dr. Mohamad Sahari Nordin, 2012) produce an influence upon the self-portraits produced. This may demonstrate how social traits could bias the self-face representation away from veridical, objective facial appearance. Notably, using the same method (Moon et al., 2020) found that self-portraits appear to be more accurate for participants who scored more highly in measures of self-esteem, explicit self-evaluation, and extraversion.

Previous research, however, has suggested that those who score highly in measures of self-esteem have a shorter gaze duration when viewing images of both self and other faces, possibly indicating a less critical evaluation of stimuli (Potthoff & Schienle, 2021). Notably, however, findings are mixed with other papers demonstrating individuals who score higher in measures of self-esteem appear to pay more attention to self-face, partner’s face, and familiar face stimuli than unfamiliar face stimuli as opposed to lower self-esteem individuals who pay, on average, the same amount of attention to all categories (Felisberti & Musholt, 2014). As such, due to the discrepancy between findings, more research is needed to establish the precise mechanism by which self-esteem interacts with self-face representation accuracy.

Furthermore, as well as demonstrating correlates between self-esteem, beliefs about one’s personality, and subjective representation, additional studies have suggested that affective traits may also bias the judgment of appearance of both the self-face and other faces. (Mirams et al., 2014) work demonstrated in their paradigm participants who scored highly in negative affect judged their appearance as a less healthy version of their own face when presented to them after the image had been manipulated in the manner suggested below; however, in individuals with high positive affect, the bias disappeared. Stimuli were classed as either healthy or unhealthy by virtue of the addition of either a red or green tinge applied, respectively.

However, concerning the spatial dimensions of the self-face representation, in terms of the distances between salient facial features, research has suggested that individuals have poor spatial knowledge of their own features, despite having no issues regarding identification. (Fuentes et al., 2013) demonstrated that within their sample, on average, participants underestimated vertical distances between features significantly, leading to an overall reduction in face height leading to a distorted face representation. Furthermore, face shape has also been reported to be a key component of self-perceived attractiveness, with studies of self-reported attractiveness in females being linked with a marked decrease in perceived facial width, and males preferring to perceive themselves as having a well-defined chin, flatter cheeks, and pronounced cheekbones. (Kanavakis et al., 2021)

Previous research has shown that in familiar faces, internal features appear to be a more reliable source of information when correctly identifying individuals than external features are, with the eyes and forehead providing the richest sources of information (Parkington & Itier, 2018). Building upon this, it has been suggested that internal mental representations of familiar faces themselves are geared more towards internal features generally (Jackson & Raymond, 2008). The inverse, however, appears to be true for unfamiliar faces, where a much greater dependence on external features is prevalent. Furthermore, this behavioural categorization appears to have a learned element, with younger children (5-6 years) appearing to focus much more heavily on external features than older children (9-10 years), who demonstrate a strong preference towards internal features for familiar faces (Bushnell, 2001). Notably however, as research into the self-face representation is in its infancy, it is not yet conclusively known if these findings translate into how the self-face is stored within memory.

More generally, it is widely accepted that face-processing is highly dependent on holistic processing. This is the view that faces are initially perceived as whole objects; this is to say, each individual facial feature is processed in parallel as opposed to the proposed piecemeal analysis of other visual stimuli (Poltoratski et al., 2021). Additional lines of research, however, argue that for the self-face specifically, processing is not purely dependent on holistic processing but also processed in a more feature-based manner (Lee et al., 2022).

Several lines of inquiry have been made into how the self-face is processed. Considering empirical evidence, it has been demonstrated in some studies that the self-face appears to exhibit a significant processing advantage over other faces (Bortolon & Raffard, 2018) with participants reacting both more quickly and with a greater degree of accuracy to images of their own face than unfamiliar faces. However, it is difficult to elicit whether this advantage is due specifically to the presentation of the self-face itself or simply due to the self-face being a highly familiar stimulus. In some works, certain personality traits have been shown to have a significant impact on reaction times, with individuals scoring highly in grandiose narcissism reacting faster to images of their own faces than controls (R. Kramer et al., 2020); again, indicating a close interaction between higher-level conceptual aspects of the self and perceptual processing of the self-face.

Electrophysiology work has also successfully demonstrated differences in the way individuals process the self-face, familiar faces, and unfamiliar faces. Variation has been demonstrated in certain event-related potentials such as the N170, with both the self-face and highly familiar face garnering enhanced N170 amplitudes (Caharel et al., 2002) as compared to unfamiliar faces, adding additional weight to the claim that familiar faces are subject to enhanced processing over unfamiliar counterparts. Notably, however, further work has successfully demonstrated the self-face specifically appears to elicit dampened P200 amplitude in comparison to other familiar faces, thus allowing for the possibility that the self-face concept may be represented as more than simply a highly familiar face (Alzueta et al., 2019). It is not possible, however, to discount the ‘familiar face advantage’ explanation entirely despite this; it stands to reason in a general sense that providing all other things remain equal, the self-face is the most familiar face to the individual in question by virtue of mere exposure alone. It remains to be seen whether the self-face is truly ‘special’ to the brain.

From a clinical perspective, the self-face is known to play a role in a wide variety of psychopathologies and as such a working model of how the self-face is represented within memory may pave the way to novel clinical applications. Disorders of self-image refer to conditions or psychological issues where individuals have distorted or negative perceptions of themselves. These disorders can significantly impact a person's mental and emotional well-being, as well as their ability to function in daily life.

There is a wide array of disorders related to abnormalities in perceptions of individual identity. For example, individuals with Body Dysmorphic Disorder (BDD) have an obsessive preoccupation with perceived flaws or defects in their physical appearance, which may be minor or even non-existent. To compensate, they may engage in compulsive behaviours such as excessive grooming, seeking reassurance, or undergoing unnecessary cosmetic procedures to try to correct these perceived flaws (Mitchell, 2017). Notably, when referring to the self-image concept in the case of BDD, it remains an open question whether these abnormalities are caused primarily by affective components such as beliefs and emotions sufferers hold towards their appearance or perceptual distortions towards their own body representations (Prnjak et al., 2022).

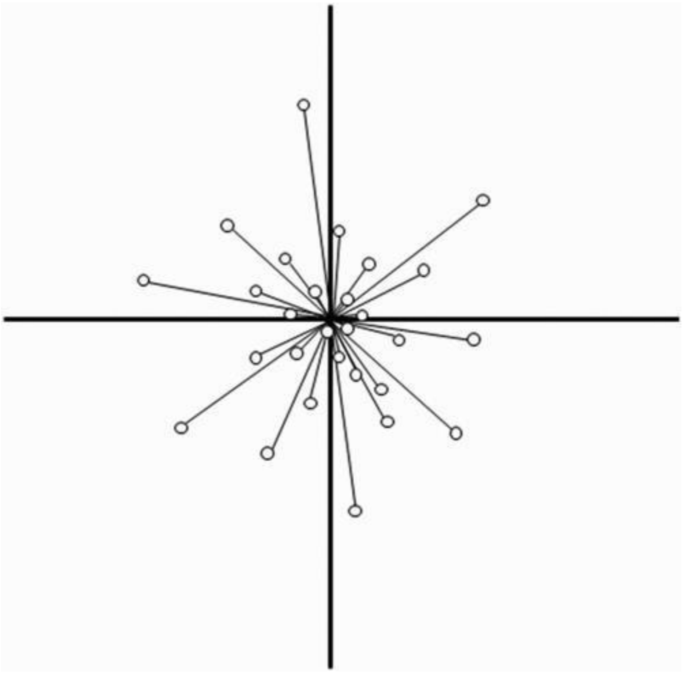
Both Anorexia Nervosa and Bulimia Nervosa are also related to a distorted self-image. Anorexia Nervosa is an eating disorder characterized by an intense fear of gaining weight and a distorted body image. Individuals with anorexia restrict their food intake to extreme levels, often resulting in severe malnutrition and other physical health complications (Hirot et al., 2016). Conversely, individuals with bulimia engage in episodes of overeating (bingeing) followed by compensatory behaviours such as self-induced vomiting, excessive exercise, or the use of laxatives. Much like individuals with anorexia, they too have a distorted self-image related to their body weight and shape (Itulua-Abumere, 2013).

Notably, distortions to the self-image also play a major role in a variety of other psychopathologies which are perhaps less immediately obvious such as social anxiety disorder and depressive disorders. Primarily characterized by intense fear and avoidance of social situations, individuals with social anxiety often have negative self-perceptions and may believe that others are scrutinizing and judging them harshly (Lee et al., 2019). Alternatively, conditions like major depressive disorder can distort an individual's self-image by leading to pervasive feelings of worthlessness, guilt, and self-criticism (Fine et al., 1993). As the self-face appears and how one chooses to identify with it appear to be so closely linked with concepts such as identity, it seems the self-face representation may be a likely candidate as a major moderating factor in many of the conditions discussed above.

One unifying theory of how pictographic representations of faces may be stored within the mind that may demonstrate where the self-face representation is stored is that of the multi-dimensional face space (Valentine et al., 2016). This model suggests that faces may be stored as individual representations within the mind, and the proximity between faces is determined by several dimensions. Known faces that are close together are proposed to be very similar to one another, whereas known faces that are a great distance apart from one another are very dissimilar. The figure below (Figure 1) depicts a visualization of how faces may be stored with a limited number of dimensions (2). Very similar faces are shown as points that are close together, whereas faces that are very dissimilar sit far apart. The origin of the face space is considered to be populated with the most ‘typical’ or average faces based on the population of faces that the individual has encountered in their lifetime. It is important to note, however, that this figure is purely for explanatory purposes, and the face space is likely to be comprised of many dimensions.

**Figure 1**

*A 2-D Representation of the Face Space.*



*Note. As depicted by (Valentine et al., 2016)*

The method to be outlined here borrows, in both a theoretical and methodological viewpoint, from that of reverse correlation; in both methods, participants select between pairs of randomly varying images based upon a given attribute. In the case of this study, participants rate how similar faces look to one another, and then how similar they look to the individual’s representation of their own face.

Notably, in both techniques individuals are never exposed to any image of their own face, and so are considered 'data-driven' and therefore relatively free from experimenter assumptions or biases (Dotsch & Todorov, 2012; Nestor et al., 2020).

Unlike reverse correlation, however, this work attempts to model Valentine's multi-dimensional face space concept (Valentine et al., 2016) in a computational manner. This approach overcomes some of the key issues presented with reverse correlation such as the production of very noisy representations (Figure 2) and gets somewhat closer to a practically applicable working model of how faces, and specifically the self-face, may be stored in the mind. It is hoped that this new technique will provide a much clearer and more vivid representation of the self-face than was previously possible from the perspective of the participant's mind's eye, and ultimately continue to add to the growing body of knowledge of how the mind works in a reproducible and falsifiable manner.

**Figure 2**

*Reconstruction Produced by Reverse Correlation.*

*Note. this image was supplied by work produced by Lara Maister (Unpublished).*

The core of the method produced here is based on earlier work by Nestor, Vettel, & Tarr (2016), an analogue of reverse correlation, who claim to have successfully reconstructed participants' long-term perceptual memories of famous faces. This approach is an ideal candidate for reconstructing the self-face representation as it is data-driven, and participants never see an image of the famous individual in question within the main experimental procedure. This methodology is, therefore, a strong candidate for also reconstructing the self-face representation as it does not allow room for participants to be biased by the act of seeing a literal photograph of themselves at any point. Overall, this means the image produced should represent a far more accurate depiction of what the participant holds within their mind's eye when asked to visualize their own face.

The main empirical method produced below largely replicates Nestor et al. (2016)'s earlier work as best possible. Major changes, however, do include the omission of certain 'spatial normalization' procedures performed as well as far greater automation of various tasks which are presumably manually achieved in Nestor et al. (2016)'s original design. Beyond these clear changes, however, the kernel of the model in question remains consistent across both Nestor et al. (2016)'s original work and this project.

This thesis will be predominantly methodological in nature, as it attempts to straddle the divide between cognitive psychology and computer science. As such, much of the work is algorithmic in nature and should be viewed as a potential 'proof of concept' as opposed to a complete solution to how the self-face is represented in memory. What follows will be the complete produced method which aims to accurately visualise the self-face representation and further empirical testing of the method via the use of an 'ideal theoretical observer' using computer vision. This will assess accuracy of the generated face representation by providing a quantifiable score between the generated representation and the real facial image of the target face. Then an additional pilot with human subjects will follow to assess human feasibility. Alternative approaches to data collection that were tested will also be presented in detail. Finally, potential implications of results, improvements, and applications are discussed.

## Method:

Experiment 1 – self-face representation image reconstruction:

In this experiment, self-face representations were reconstructed using the adapted method from (Chang et al., 2017), based on (i) artificial data from an 'ideal theoretical observer' model utilizing the Openface algorithm (Amos et al., 2016), and (ii) human data from 18 recruited participants. Both results are assessed using a number of mathematical validation tests and compared.

Participants:

Participants (N=18) were recruited using both the SONA recruitment platform and word of mouth. All participants completed pre-screening for eligibility in the form of the Vividness of Visual Imagery questionnaire (VVIQ-2) (Campos, 2011) (For data and codes associated with this work, please see appendix) prior to testing to ensure they had no notable deficits in visual memory. All participants had normal or corrected to normal vision and no known history of visual/neurological disorders. Participants ages ranged between 18 – 35 years of age and came from a variety of ethnic backgrounds, namely, Indian (2), Black (1), Iranian (2), White (12) and Chinese (1) with stimuli tailor made to match each ethnic group.

Stimuli:

Fifty-seven face images were selected as stimuli from the appropriate database, selected to match each individual's participant's ethnic background as per (Chang et al., 2017)'s paradigm. To the best of the experimenter's knowledge, these were unfamiliar to participants. Stimuli were gathered from several free-to-use databases, namely: Radbound Faces Database (Langner et al., 2010), Oslo Face database (https://sirileknes.com/oslo-face-database/ (No

citation information available)). Chicago Face Database (Ma et al., 2015). Stimuli were selected based upon several criteria ensuring all faces selected fell between the ages of 18 – 35 years of age and the following traits: ethnicity, attractiveness, dominance, trustworthiness, and unusualness were controlled for via exclusion of any images which fell beyond +/- 2 standard deviations from the mean on any of the traits using the codebooks provided by the face databases (For data and codes associated with this work, please see appendix). All stimuli were then placed on a transparent background and oval-cropped around the jaw and hairline ensuring peripheral visual features were removed. A consistent 8-pixel feather was also applied to the cropped edge. Stimuli were then colour normalized by equating the mean values of each CIEL*a*b colour channel across each demographic (For data and codes associated with this work, please see appendix) Stimuli were then placed on a 380*570 canvas with a black background. Finally, each stimulus was resized to 105*158 pixels ready for display. All stimuli were forward facing, looking directly into the camera and displayed a neutral expression. Immediately obvious artefacts such as glasses, heavy-make-up, long beards and jewellery were also controlled for so as not to be included within the stimuli set (see figure 3).

**Figure 3**

*White Female Stimuli Used in the Above Paradigm.*

*Note. Image available from, Oslo Face database (https://sirileknes.com/oslo-face-database/ (Nocitation information available)*

**Experimental procedures:**

All participants gave full informed consent prior to testing for both experimental sessions (each session being ~1 hour in length) which occurred over two days. All procedures were carried out in accordance with Bangor Universities Research Ethics Guidelines and were approved by the Bangor Universities Research Ethics Board. The experiment consisted of a face-space estimation task, and a self-face task. Both experimental tasks were developed in MATLAB R2021b using Psychtoolbox-3. Participants first completed the Face-space estimation task in its entirety before proceeding to the self- face task. Prior to completing either task, a photograph was taken of each participant displaying a neutral expression following a set of pre-defined guidelines (For data and codes associated with this work, please see appendix).

Face-space estimation task:

In brief, following a 500ms fixation each trial the Face-space estimation task consisted of participants being presented with two faces side by side for a total of 2000ms and then being required to provide a rating based on the similarity of the previous two faces presented, further details of the task are provided below. Counterbalancing was used in the Face-space estimation task to ensure each face was paired with each other face exactly once, resulting in 1596 trials in total. This task was split over multiple sessions ensuring participants did 798 trials per session. Participants were given breaks every 100 trials.

Stimuli were never paired with themselves and in total over 1596 trials each face appeared 28 times on the left and 28 times on the right in line with the task proposed by (Chang et al., 2017) although different stimuli were used.

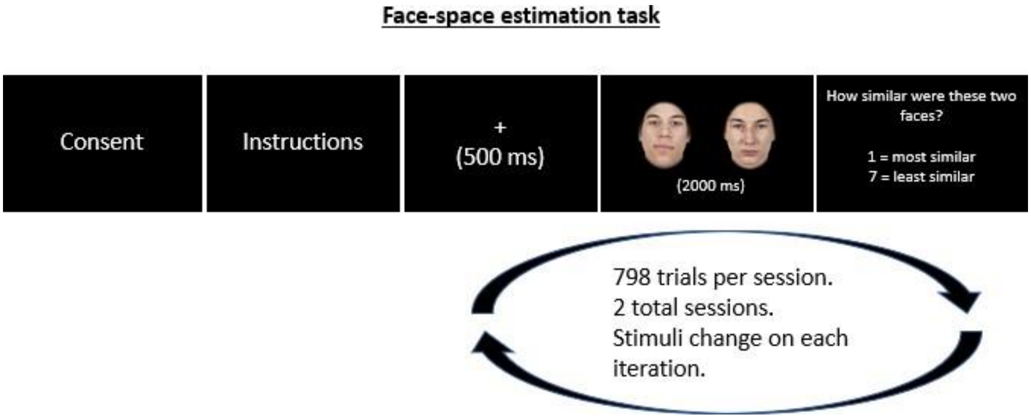
After providing consent and confirming to have read and understood the onscreen instructions, participants saw two faces presented side by side which they needed to provide a similarity score for between 1 – 7 by pressing a number key for each face pair with 1 indicating a large similarity between both faces and 7 indicating a large dissimilarity. Stimuli pairs appeared immediately after a fixation cross in the centre of the screen with a 500ms duration. Stimuli pairs remained onscreen for 2000ms at which point participants were prompted to give there rating and once again reminded of the scale. Keystrokes administered prior the prompt where not recorded and participants were unable to proceed to the next trial under a rating was given. Participants had an unlimited time to provide each response.

All participants were sat directly in front of the monitor and used a chin rest for the duration of the task. Each individual face stimuli subtended a visual angle of 2.6°

× 4° from a viewing distance of 60 cm and was displaced 2.4° from the centre of the screen. (See Figure 4)

**Figure 4**

*Procedural Workflow for Face-Space Estimation Task.*




Self-face task:

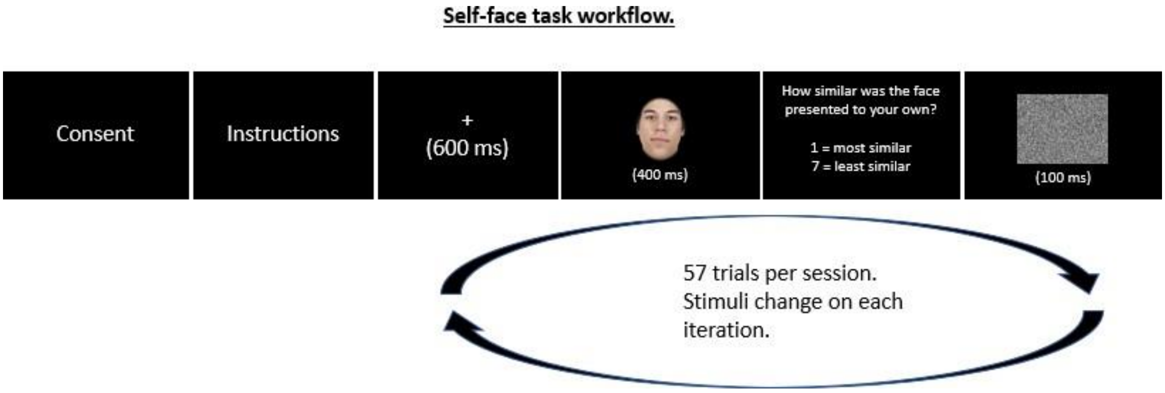
The self-face task was ~2 minutes in duration and consisted of 57 trials. Participants completed this task immediately after completing the Face-space estimation task.

After providing consent and confirming to have read and understood the onscreen instructions participants again provided a similarity score between 1 – 7 using the same scale as the Face-space estimation task (1 = most similar, 7 = least similar). In this task, each of the 57 stimuli were presented individually once in the centre of the screen, immediately following a central fixation cross with a duration of 600ms. Stimuli remained onscreen for 400ms at which point participants where then prompted to give there rating. However, unlike the Face-space estimation task, in this task participants where asked ‘How similar was the face presented to your own?’

and once again reminded of the scale. Keystrokes administered prior the prompt where not recorded and participants were unable to proceed to the next trial until a rating was given. Immediately after each rating a white noise mask appeared in the centre of the screen for 100ms before proceeding to the next trial. (See Figure 5)

**Figure 5**

*Procedural Workflow for Face-Space Estimation Task.*




All participants were sat directly in front of the monitor and used a chin rest for the duration of the task with stimuli subtending an angle of 2.6° × 4° from a viewing distance of 60 cm. For all tasks (both memory and perception based) trial order was randomized, and stimuli were presented on a black background.

Artificial data collection:

Artificial data was gathered via the use of the OpenFace convolutional neural network (Amos et al., 2016) which produces comparative scores between the dissimilarity of two images of faces based upon the discrepancies in L2 distance between fiducial points between images. Initially, all stimuli for white male participants (N=57) were fed into the network and compared against one another, this theoretically gave the most objective ‘face-space estimation’ available. Sixty photographs of faces of white men (for data and codes associated with this work, please see appendix) were then selected and each image was then compared against each of the fifty-seven stimuli faces. Data was then arranged and processed to mimic that of human participant data.

Reconstruction procedure:

The reconstruction procedure was based upon the earlier work of (Chang et al., 2017). For each individual participant, including ‘artificial’ OpenFace participants (Amos et al., 2016), faces were reconstructed separately utilizing MATLAB R2021b and R 4.1.2. The procedure consisted of several steps which will be outlined below.

For face space estimation-based data, firstly similarity scores from the fifty-seven faces displayed in the Face- space estimation task was arranged into a symmetrical dissimilarity matrix (57*57) with each row and column representing one face. Therefore, each cell represented the perceived dissimilarity between each possible pair of faces. As is usual in dissimilarity matrices, the diagonal contained zeros, reflecting an assumed null level of dissimilarity between identical faces.

Classical Metric Multidimensional Scaling was then applied to the dissimilarity matrix with a maximum of twenty dimensions as per the earlier well validated work of (Chang et al., 2017). This procedure created the coefficients required to represent a multidimensional face space for each participant. For each participant, their memory-based scores were then aligned with their multidimensional face spaces using Procrustes alignment, this was achieved by initially generating an additional multidimensional face space including the self-face and then aligning both face spaces to result in one final face space which contained the self-faces placement within the pre-existing face-space garnered from the earlier ‘face-space estimation’ task. This stage of analysis was completed in R and resulted in two independent sets of coefficients; one reflecting the face-space based on the fifty-seven facial stimuli provided, and one additionally containing the location of the self-face within this face-space. The self- face coefficients were z-scored after alignment as per (Chang et al., 2017), procedure to ensure the same scale was applied to all coefficients throughout.

The next stage of analysis was performed in MATLAB. Initially, an overall average face image was produced which consisted of a non-weighted average of all fifty-seven of the faces displayed during the experiment (the stimuli). This image served as the base, to which classification images, which are fragments of the complete facial images provided by the original stimuli containing potentially identifying information such as eyebrow densities. These images are then overlaid on to the base ‘average’ face. Coefficients were then sorted into positive and negative values for each dimension separately, and then each were combined resulting in two separate average images per dimension based upon either the weighted average positive or negative coefficients.

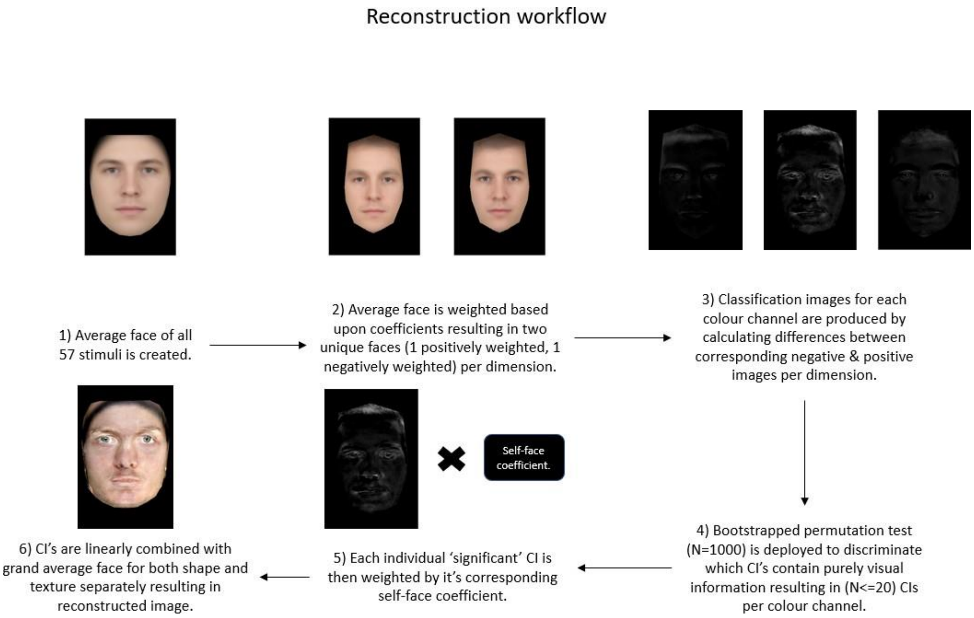
These weighted average images serve as the blueprint for which classification images for each dimension can be derived. Classification images (CI) for each dimension were created by calculating the difference between the corresponding positive and negative images for each dimension. This process results in CI’s which encapsulated the specific visual features associated with each of the 20 dimensions of the derived face space. Of note, it is possible that not all dimensions encode purely visual information using this approach, as similarity is a multifaceted concept which spans beyond the visual domain which could skew results significantly. Therefore, a bootstrapped permutation test was deployed to assess which CI's contained significant visual information using a pixel-wise approach over 10 random permutations in which a selection of randomly generated CI's were tested against chance against the observed dimension provided by the participants' similarity scores. This was performed by assessing CI significance by randomizing images with respect to their co-ordinates on each dimension, and then recomputing CIs for a total of 10 permutations.

This process ultimately results in several CI's (N<=20 per channel) that characterize distinct, salient visual dimensions of the face space, for both shape and texture separately. Of note, 'shape' files (x y co-ordinates) were extracted using an automated landmarking tool ('dlib'), (Anonim, 2019) and are a required dependency for each stimulus face to proceed with the reconstruction procedure.

After 'significant' CI's are extracted, each individual CI is then weighted by its corresponding self- face coefficient, indicating where the target (self-face) is in the modelled multi-dimensional face space. Finally, the now correctly weighted CIs are linearly combined with the grand average face (for shape and texture separately). Shape and texture vectors are then combined, resulting in the finished reconstructed image (see Figure 6).

**Figure 6**

*Reconstruction Procedure.*




*Note. The image created in 'step 2' here was produced using the 'InterFace' toolbox, this software or any associated codes however were not used anywhere else in this work. (R. S. S. Kramer et al., 2017)*

## Results (Experiment 1):

Analysis plan:

Analysis took a three-pronged approach; firstly, reconstruction accuracy was assessed using the OpenFace computational neural network (Amos et al., 2016) which aims to provide one dissimilarity score between each reconstruction and each participants' real face. In brief,

> "OpenFace… outputs the predicted similarity score of two faces by computing the squared L2 distance between their representations. A lower score indicates two faces are more likely of the same person. Since the representations are on the unit hypersphere, the scores range from 0 (the same picture) to 4.0".

This allowed for the most direct, objective one-to-one comparison between reconstructions and photographs.

Following OpenFace analysis (Amos et al., 2016), images were then split into their component parts (shape and texture) for inspection. The Euclidian/Pythagorean distance between reconstruction and photograph was calculated for both shape and texture separately. Shape, in this context refers to the 81 landmark points, in the form of [x, y] co-ordinates, which are plotted on all faces via the dlib landmarking engine (Anonim, 2019). Texture refers to the individual pixels which comprise each image, reflecting colour and luminance variations. Importantly, before Euclidean distances were calculated, images were first Procrustes aligned so that major image features between images overlapped as best possible before attempting to compare them.

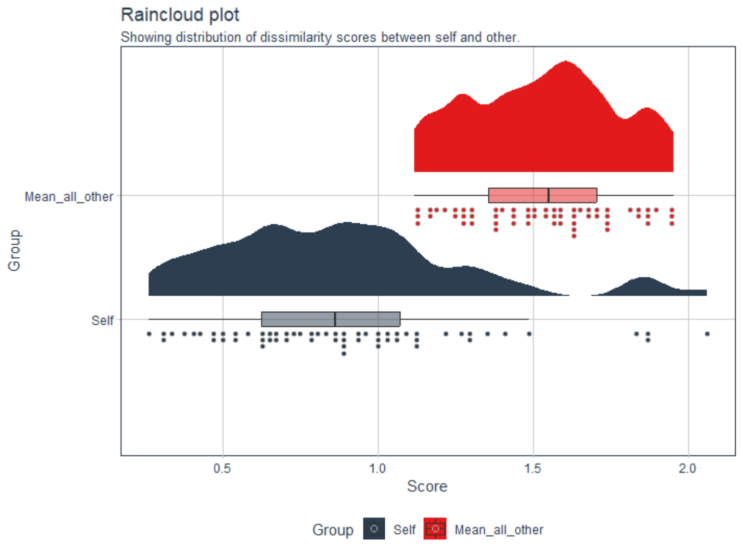
Results – Ideal theoretical observer data:

A paired samples t-test was deployed to ascertain if reconstructed images contained identifying information. This was achieved by computing cross-individual comparisons between reconstructed faces and all other faces in the sample based upon OpenFace (Amos et al., 2016) scores which were ascertained as outlined above. Specifically, the mean ‘distance’ of each reconstruction to its corresponding target face (self) as provided by OpenFace (Amos et al., 2016) dissimilarity score was compared with the mean distance against all other faces which were not the target face within the sample (non-self).

The results of the t-test were significant; self: $M = 0.89$, $SD = 0.40$; non-self: $M = 1.53$; $SD = 0.24$, 95% CI for the mean difference = [−0.74, −0.55]), paired-samples $t(59) = -13.72$, $p < .001$, Cohen’s $d = 1.95$. This result demonstrates that the self-face reconstructions were significantly less dissimilar, i.e., more similar, to the self-face targets than any other (non-self) targets as assessed by OpenFace (Amos et al., 2016), (see Figure 7)

**Figure 7**

*Cross-Individual Comparisons Between Reconstructed Faces and all Other Faces in the Sample Based Upon OpenFace Scores.*




*Note. N = 60. Data from this test was retrieved from the OpenFace convolutional neural network (Amos et al., 2016) which provides dissimilarity scores between any two face images. Dots signify individual datapoints, box and whisker plots signify mean values of each group and upper curves reflect density distribution of datapoints.*

Reconstruction accuracy was also assessed by calculating similarity between reconstructed and target (real) faces using Procrustes analysis and the calculation of Euclidian distances between sets of points broadly in line with (Chang et al., 2017) approach. For shape, this approach entailed comparing each reconstructed face with its target by firstly landmarking both the target faces and the reconstructed faces using the 'dlib' facial landmarking engine (Anonim, 2019). Following this, the reconstructed landmark points where then Procrustes aligned with the target landmark set and then also mapped onto the target face (See Figure 8). Euclidian distances between the target and reconstruction set were then computed.

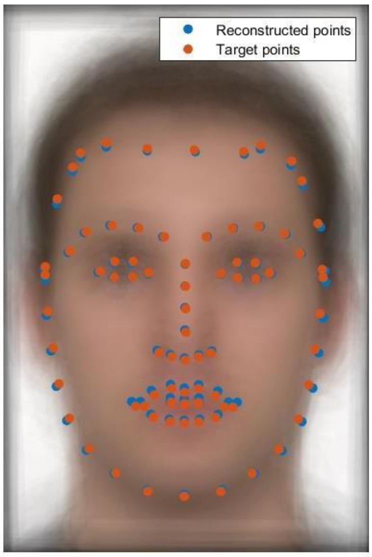
A Paired samples t-test was used to ascertain whether the Euclidean distances between all self vs. target landmark point comparisons and all self vs. non-target landmark point comparisons significantly differed. This resulted in self: M = 7.23, SD =2.74; non-self: M = 8.04; SD =1.72, 95% CI for the mean difference = [−1.46, −0.16], paired-samples t (59) = −2.49, p < .01, Cohens d = 0.32. This demonstrates that the mean Euclidean distance between reconstructed points and their target counterparts is significantly lower between reconstruction and target, than between reconstruction and all other possible non-self-targets. (See Figure 9)

Following this, the Euclidian distance for texture was calculated on a pixel-by-pixel basis as a product of Procrustes aligning each image, then vectorizing it and taking the Pythagorean distance between both images. A paired samples t-test was then performed between the mean of all self vs. target Euclidian distance scores following Procrustes alignment and the mean of all reconstructed self vs. non-self-target Euclidian distance scores. This resulted in self: M = 522.31, SD = 35.40 non-self: M

= 524.46 SD = 4.10, 95% CI for the mean difference = [ -11.71, 7.41], paired-samples t (59) = -0.45, p = 0.65. This demonstrates that the mean Euclidean distance between reconstructed faces and their target counterparts is not significantly lower between reconstruction and self-target, than between reconstruction and all other possible non-self-targets when comparing pixel colour and luminance values

**Figure 8**

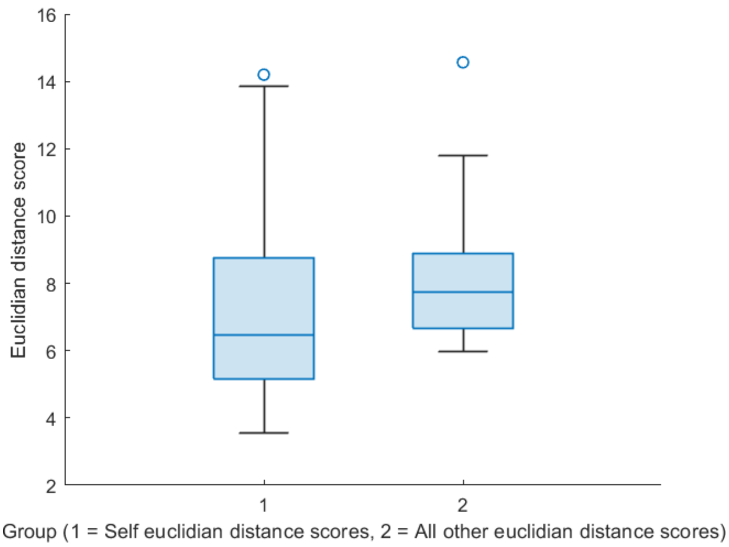
*Reconstructed Landmark Points Procrustes Aligned With the Target Landmark Set and Mapped Onto the Target Face.*




*Note. point sets and face depicted here are an overall average of all (N = 60) participants from artificial data and their respective landmark points.*

**Figure 9**

*Euclidean Distances Between all Self vs. Target Landmark Point Comparisons and all Self vs. Non-Self-Target Landmark Points*




*Note. N = 60, Figure produced in MATLAB R2021b. Boxes represent median values at the point of the line and interquartile ranges for each group either side of the line (lower & upper quartile), whiskers represent minimum and maximum range indicating spread of scores. Blue circles are outliers.*

Results – Human Participant data:

The same analyses were carried out on the data derived from the human participants. First, a paired samples t-test was deployed to ascertain if reconstructed images contained identifying information. Unfortunately, this could only be carried out within a group of participants that were homogenous in terms of gender and ethnicity. This was achieved by computing cross-individual comparisons between reconstructed faces and all other faces in the sample for white female participants, which proved to be the largest overall group (N = 8). As before, data from this test was derived from the OpenFace convolutional neural network (Amos et al., 2016) which provides dissimilarity scores between any two face images. Specifically, the mean ‘distance’ of each reconstruction to its corresponding target face (self) as provided by OpenFace dissimilarity score was compared with the mean distance against all other faces which were not the target face within the sample (non-self) (see Figure 10).

The results of the t-test were non-significant (self: M = 1.31, SD = 0.40; non-self: M = 1.37 SD = 0.18, 95% CI for the mean difference = [−0.52, 0.40]), paired-samples t (7) = −0.31, p = 0.76. This result demonstrates that the self-face reconstructions were not significantly more similar to the self- face targets than any other (non-self) targets as assessed by OpenFace (Amos et al., 2016)

**Figure 10**

*Cross-individual Comparisons Between Reconstructed Faces and All Other Faces in the Sample for White Female Participants.*

*Note. N = 8. Data from this test was retrieved from the OpenFace convolutional neural network (Amos et al., 2016) which provides dissimilarity scores between any two face images. Dots represent individual data points. Boxes represent median values at the point of the line and interquartile ranges for each group either side of the line (lower & upper quartile), whiskers represent minimum and maximum range indicating spread of scores.*

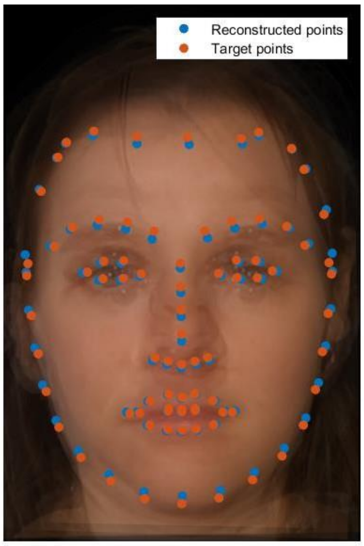
Reconstruction accuracy for participant data was also assessed by calculating similarity between reconstructed and target faces using Procrustes analysis and the calculation of Euclidian distances between sets of points broadly in line with (Nestor et al., 2016) approach. For shape, this approach entailed comparing each reconstructed face with its target by firstly landmarking the target face using the ‘dlib’ facial landmarking engine (Anonim, 2019). Following this, the reconstructed landmark points were then Procrustes aligned with the target landmark set, and then also mapped onto the target face (See Figure 11). Euclidian distances between the target and reconstruction set were then computed.

A Paired samples t-test was used to ascertain whether the Euclidean distances between all self vs. target landmark point comparisons and all self vs. non-self-target landmark point comparisons significantly differed. This resulted in:

self: $M = 9.2$, $SD = 2.4$; non-self: $M = 9.5$ $SD = 1.6$, 95% CI for the mean difference = [−2.09, 1.63], paired-samples $t(7) = -0.30$, $p = 0.78$. This demonstrates that the mean Euclidean distance between reconstructed points and their target counterparts is not significantly lower between reconstruction and self-target, than between reconstruction and all other possible non-self-targets (see Figure 12).

**Figure 11**

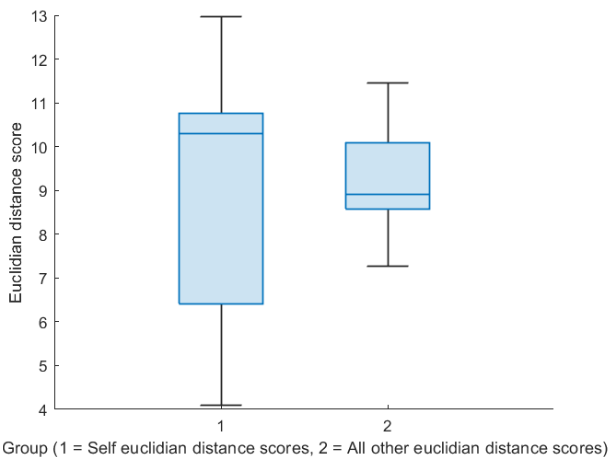
*Reconstructed Landmark Points Procrustes Aligned with the Target Landmark Set and Mapped Onto the Target Face.*




*Note. point sets and face depicted here are an overall average of all (N = 8) participants and their respective landmark points.*

**Figure 12**

*Euclidean Distances Between all Self vs. Target Landmark Point Comparisons and all Self vs. Non-Self-Target Landmark Points.*



*Note. N = 8, Figure produced in MATLAB R2021b.*

Following this, the Euclidian distance for texture was calculated on a pixel-by-pixel basis as a product of Procrustes aligning each image, then vectorizing it and taking the Pythagorean distance between both images. A paired samples t-test was then performed between the mean of all self vs. target Euclidian distance scores following Procrustes alignment and the mean of all self vs. non-self-target Euclidian distance scores. This resulted in:

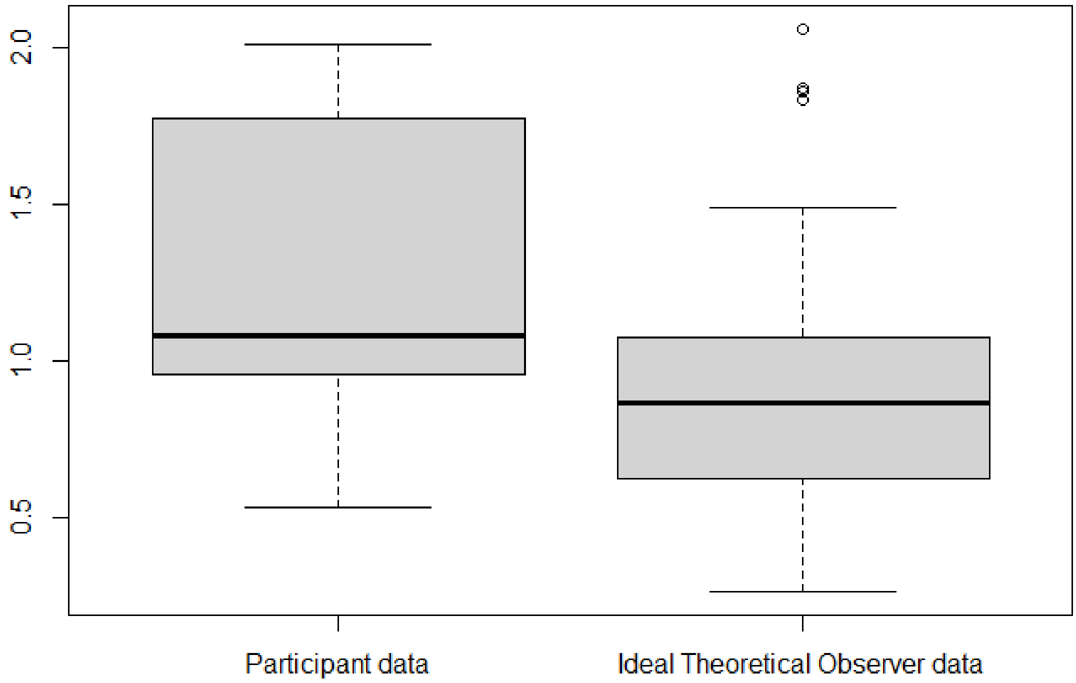
self: $M = 345.73$, $SD = 32.25$ non-self: $M = 348.83$ $SD = 13.31$, 95% CI for the mean difference = [- - 29.15, 22.96], paired-samples $t(7) = -0.28$, $p = 0.79$. This demonstrates that the mean Euclidean distance between reconstructed faces and their target counterparts is not significantly lower between reconstruction and self-target, than between reconstruction and all other possible non-self-targets.

A Welch's t-test was deployed to ascertain if there was a significant difference between all collected human participant data (Earlier analysis focused upon the largest group only as many demographics contained only 1 participant) and ideal theoretical observer data. This was achieved by computing dissimilarity scores between reconstructed images and their targets using the OpenFace CNN (Amos et al., 2016) for both participant data and Ideal theoretical observer data. Specifically, this tested the mean dissimilarity scores of all self-reconstructions to their target faces as gathered by participant data ($n=18$) against all self- reconstructions to their target faces as generated by the use of the ideal theoretical observer.

(Participant data: $M = 1.29$, $SD = 0.45$; Ideal theoretical observer data: $M = 0.89$ $SD = 0.40$, 95% CI for the mean difference = [0.16, 0.65]), Welch's $t(25.7) = 3.39$, $p = 0.02$. This result demonstrates that there was a significant difference between Ideal theoretical observer data and participant data as assessed by OpenFace (Amos et al., 2016) with artificial data surpassing participant data by producing significantly lower dissimilarity scores. (See Figure 13)

**Figure 13**

*Boxplot Showing Mean Dissimilarity Scores Between Participant and Ideal Theoretical Observer Data.*




*Note. Participant data (N = 18), Ideal theoretical observer data (N = 60). This result demonstrates that there was a significant difference between Ideal theoretical observer data and participant data.*

Following this, a further Welch's t-test was deployed to ascertain if there were significant differences between participant and OpenFace (Amos et al., 2016) shape and texture data. This was achieved by comparing participant self-target vs non-self-target distances for both shape and texture separately against OpenFace (Amos et al., 2016) self-target vs non-self-target differences. This resulted in the following:

Shape:

(Participant data: M = 9.0, SD = 3.0; Ideal theoretical observer data: M = 7.23 SD = 2.74, 95% CI for the mean difference = [0.782, 4.32]), Welch's t (8.62) = 1.58, p = 0.15 This result demonstrates that there was not a significant difference between Ideal theoretical observer data and participant data as assessed Euclidian distance landmark point scores.

Texture:

(Participant data: M = 345.73, SD = 32.25; Ideal theoretical observer data: M = 522.31 SD = 35.40, 95% CI for the mean difference = [ -204.18, -148.97]), Welch's t (9.40) = - 14.38, p = 0.01. This result demonstrates that there was a significant difference between Ideal theoretical observer data and participant data as assessed pixelwise Euclidian distance scores.

## Experiment 2 – Alternative approaches to data collection and test-retest validity analysis:

Experiment 1 demonstrated that, although the task worked well for our ideal theoretical observers, results were mixed for our human participants. Concerns about time taken and motivation of participants during the face-space estimation task led to attempts to shorten the task time. Due to the fact that the standard face-space estimation task, adopted from Chang et al. (2017), was based on pairwise comparisons of 57 facial identities, this required 1596 trials in total which was a heavy testing burden. One possible solution may be the multi-arrangement method

(Kriegeskorte & Mur, 2012) demonstrated below, which involves presenting an array of faces on each trial, drastically reducing the number of trials required. This task has been shown to perform well (Verheyen et al., 2022). In the following experiment, a multi-arrangement method of estimating the face space was compared to the traditional pairwise method, with a view to its potential adoption for the self-face reconstruction task.

Participants:

Participants (N=39) were recruited using both the SONA recruitment platform and word of mouth. All participants had normal or corrected-to-normal vision and no known history of visual/neurological disorders. Participants' ages ranged between 18 – 35 years of age and came from a variety of ethnic backgrounds with stimuli tailor-made to match ethnic groups as best possible. All participants gave full informed consent prior to testing for both experimental sessions. All procedures were carried out per Bangor Universities Research Ethics Guidelines and were approved by the Bangor Universities Research Ethics Board.

Stimuli:

For each demographic 24 face images were selected as stimuli as this number was as close as possible to getting participants to do half of the original task (57 stimuli) for piloting purposes, which, to the best of the experimenter's knowledge were unfamiliar to participants. Stimuli were gathered from several free-to-use databases, namely: Radboud Faces Database, Oslo Face Database, and Chicago Face Database (Ma et al., 2015). Stimuli were selected based upon several criteria ensuring all faces selected fell between the ages of 18 – 35 years of age and the following traits: ethnicity, attractiveness, dominance, trustworthiness & unusualness were controlled for using the codebooks provided by the face databases by calculating the mean and standard deviation for each trait and then removing any faces from the stimuli set which fell +/- 2 standard deviations from the mean.

All stimuli subjects were then placed on a transparent background and then oval cropped ensuring only key visual information remained. A consistent 8-pixel feather was also applied to each face. Stimuli were then colour normalized by equating the mean values of each CIEL*a*b colour channel across each demographic (for data and codes associated with this work, please see appendix). Stimuli were then placed on a 380*570 canvas with a black background.

Finally, each stimulus was resized to 105*158 pixels ready for display. All stimuli were forward- facing, looking directly into the camera, and displaying a neutral expression. Immediately obvious artefacts such as glasses, heavy make-up, long beards, and jewellery were also controlled for.

Notably, during multi-arrangement tasks, though the same images were used, the size at which they were displayed differed due to screen size and resolution restrictions which accompany displaying multiple images on the same monitor as opposed to just two. However, all images remained visible and discernible from each other on a trial-by-trial basis. For the pairwise task, the visual angle used mimicked that of the initial face-space estimation task.

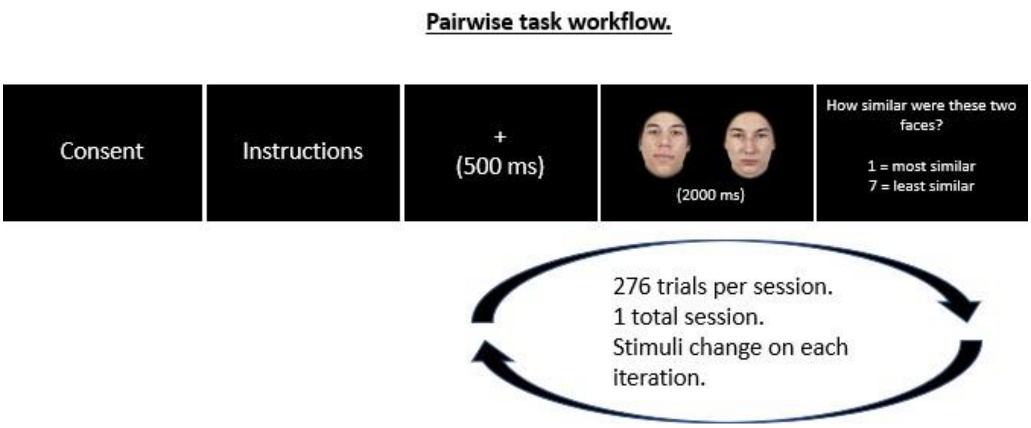
Notably, the stimuli used mirrored those used in the initial face-space estimation task and were also therefore subjected to the same pre-selection requirements.

Procedure (pairwise):

All participants sat directly in front of the monitor and used a chin rest for the duration of the task. Each face stimuli subtended an angle of 2.6° × 4° from 60 cm and was displaced 2.4° from the centre of the screen. (See Figure 14 for procedural task workflow) Notably, this task matches that of the earlier pairwise method regarding procedure but with 24 stimuli as opposed to the 57 used in the ‘face space estimation’ task.

**Figure 14**

*Pairwise Task Procedure*




Procedure (Multi-arrangement).

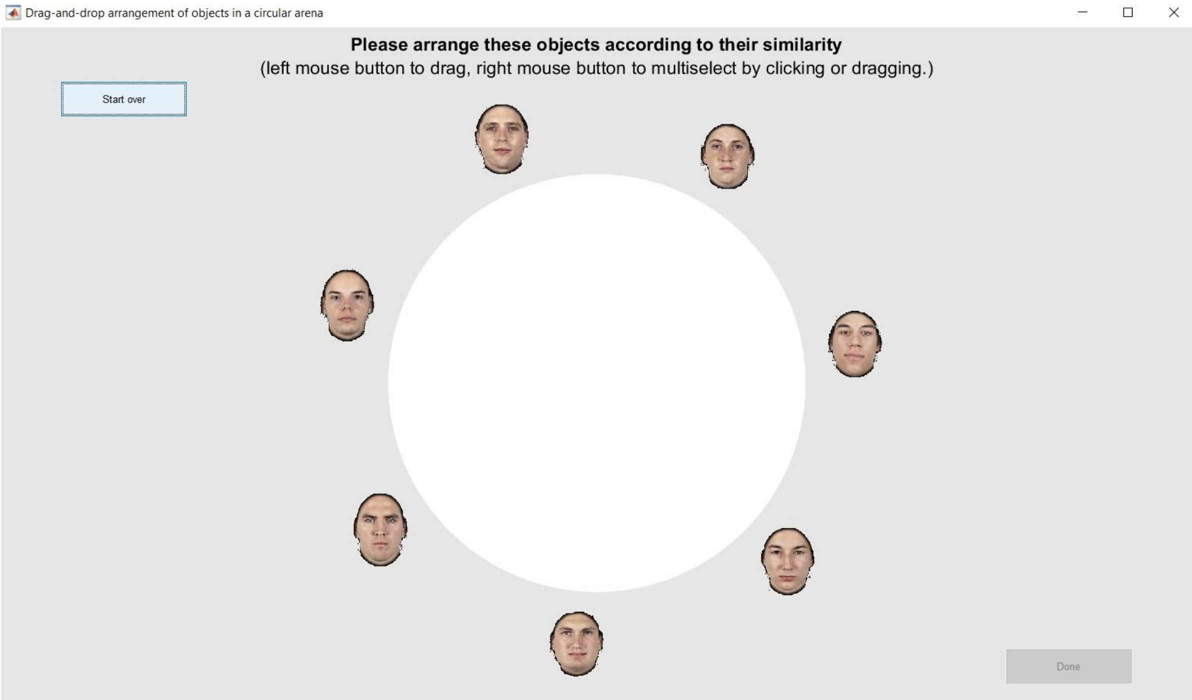
This task was programmed using MATLAB R2021b after being adapted from a task developed by (Kriegeskorte & Mur, 2012).

After providing consent and confirming to have read and understood the onscreen instructions participants were required to move a selection of seven faces onscreen via drag and drop operations using a mouse in each trial. The objective was to place images which they believed to be perceptually similar to one another close together and images which they believed to be dissimilar further apart from one another within the pre-defined arena, which was represented as a large white dot in the middle of the screen. Initially, all 7 images which composed each trial were circularly seated around the arena with no images inside the arena. On each trial participants had the option to start (the trial) again should they wish by pressing ‘Start over’ located on the top right. Should participants press this, stimuli positions were randomized once more as they returned to the original seating confinements (i.e., outside of the arena). Participants were unable to proceed to the next trial until all stimuli resided within the arena at which point, they were able to press ‘Done’ located in the lower left of the screen to proceed. Please see ‘Figure 15’ and ‘Figure 16’ for an example image of one trial both pre and post-completion. Participants were also able to zoom in at any time by pressing either windows + or – accordingly. A total of 7 images were presented to participants on each trial.

A Steiner system of order C (24,7,2) (Ballico et al., 2021) was deployed producing a set trial order, meaning a fixed trial count of 17 trials was achieved whilst only ever displaying 7 images onscreen at any given time. Similarity ratings were garnered for this task by taking the Euclidian distance between all stimuli placed by participants within the pre-defined arena ultimately ending with the same number of similarity ratings as the pairwise task (276).

***Figure 15***

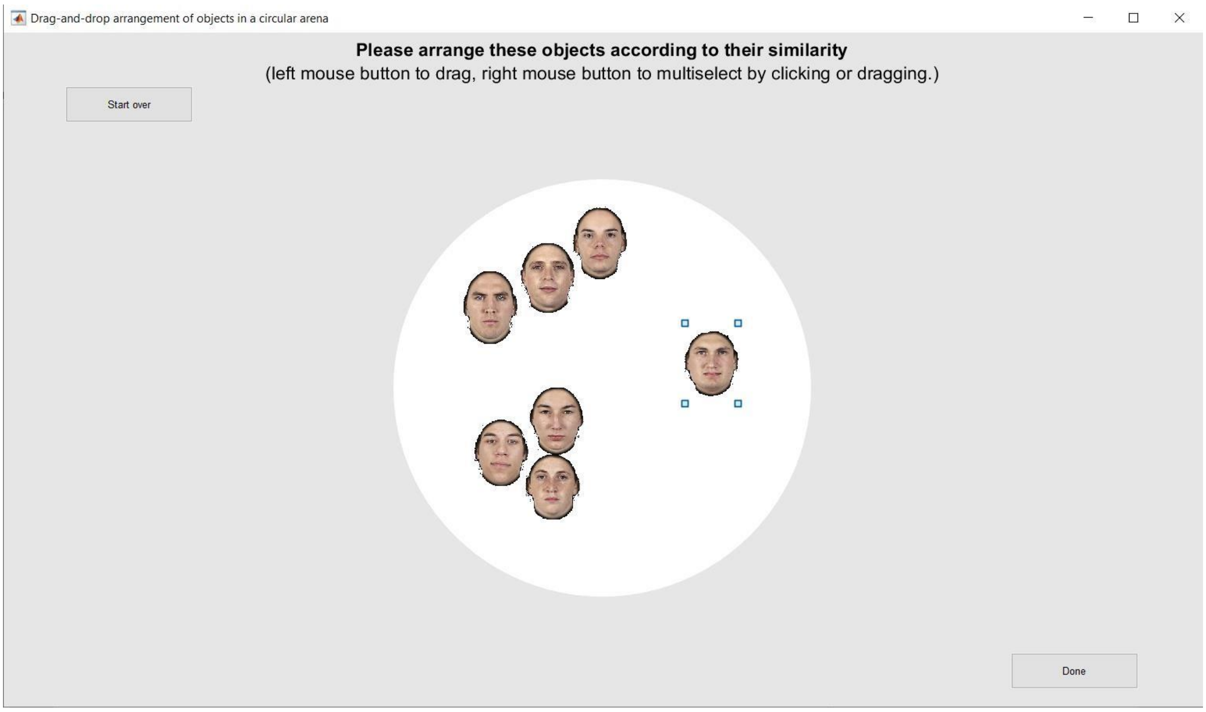
*Example Multi-Arrangement Starting Trial.*




*Note: Task produced in MATLAB R2021b.*

**Figure 16**

*Example Multi-Arrangement Completed Trial.*




*Note: Task produced in MATLAB R2021b.*

Design(s):

Counterbalancing was used in the pairwise task (See Experiment 1 for further details) to ensure each face was paired with every other face exactly once, resulting in 276 trials in total. Participants were given breaks every 100 trials. Stimuli were never paired with themselves and in total, over 276 trials each face appeared either 11 times on the left and 12 times on the right or vice versa.

During the multi-arrangement task images were displayed based upon the Steiner system (Ballico et al., 2021) in sets of 7 exclusively.

Experimental procedures:

The experiment consisted of several tasks requiring participants to provide similarity ratings between a set of faces, the sessions took place over 2 days with sessions lasting ~ 30 minutes per session and one session being performed each day. All experimental tasks were developed in MATLAB R2021b, the pairwise task specifically was developed using Psychtoolbox-3. The sessions consisted of participants completing both the multi-arrangement and pairwise sections of the task back-to-back, however the order of tasks was counterbalanced across participants, so the participants could not complete the pairwise task followed immediately by the multi-arrangement task or vice versa across both testing days.

Test-retest protocol:

Participants were required to perform each task a grand total of two times each over a period of 3 days in the same conditions following the same procedure. I.e., each participant performed both the multi- arrangement and pairwise tasks twice.

## Results (Experiment 2):

Spearman's rank correlation was computed to assess the relationship between pairwise and multi- arrangement tasks for male and female participants separately. Specifically, this aimed to assess if similarity judgements provided by males and females across both tasks (pairwise, multi-arrangement) are comparable. Scores were computed as an average of the similarity ratings each participant had provided over both the test days they attended.

There was a very weak negative non-significant correlation between the two tasks for female participants, r (28) = - 0.01, p = .87 over averaged gender matrices of both tasks. There was a very weak positive non-significant correlation for male participants. r (11) = 0.06, p = 0.30 again over averaged gender matrices of both tasks. Notably, combining both sexes together also failed to produce any significant results. r (39) = 0.03, p = 0.48.

Finally, 10 additional participants were recruited to perform the pairwise task twice over a period of two days after fulfilling the same selection criteria outlined above to assess the test-retest validity of the pairwise protocol. As the below table demonstrates (Table 1), the results of the same participant scores, doing the same task in the same conditions on a different day vary significantly with only one participant indicating a moderate significant positive correlation (r = 0.5, p = 0.01) between both tasks:

**Table 1**

*Test-Retest Correlations of Participants Performing the Pairwise Task Protocol.*

| Participant: | Spearman's Rho: | P value: |
|---|---|---|

| | | |
|---|---|---|
| 1 | -0.02075 | 0.731456 |
| 2 | 0.017081 | 0.777561 |
| 3 | 0.029599 | 0.624415 |
| 4 | 0.544123 | 0.000001 |
| 5 | -0.03103 | 0.607693 |
| 6 | 0.042568 | 0.481243 |
| 7 | -0.0034 | 0.955098 |
| 8 | 0.163466 | 0.006495 |
| 9 | 0.006294 | 0.917099 |
| 10 | 0.324749 | 0.000001 |

## Discussion:

This project aimed to produce a method of visualizing long-term perceptual memories of faces to visualize the self-face representation. It aimed to produce the method, test it via the use of an 'ideal theoretical observer' (OpenFace) (Amos et al., 2016) as an initial proof of concept and then pilot with human subjects to assess feasibility. With the ideal observer data, results were promising with self-face reconstructions being significantly more similar to self-face targets than non-self-targets.

Unfortunately, this finding was not mimicked in participant data leading to alternative approaches to data collection being trialled as well as test-retest validity measures to be run. Both initial test-retest of the original method of pairwise data collection as well as well as alternative approach (multi-arrangement) analysis produced only poor non-significant correlations within participants. In the following section, potential implications of results, improvements, and applications will be discussed.

One potential limitation of this approach pertains to the 'face-space estimation' task which sits at the core of the reconstruction approach. To produce the estimate, pairwise ratings are required across the entire stimulus set. Using the current approach this means participants are required to complete n (n- 1)/2 possible trials with 'n' relating to the number of items in the set (57). Overall, this means participants are required to sit and produce pairwise ratings for several hours in total. This could potentially lead to extraneous factors such as boredom and inattention being a key factor in the validity of results garnered. One potential venue for further exploration which may limit these confounds however could simply involve splitting the 'face-space estimation' task into multiple shorter sessions providing the task (face-space estimation) exhibited test-retest validity although initial results here of this shortened version evidently do not look promising.

Furthermore, neither of the tasks is particularly feature specific as they do not display any facial features in isolation but rather present a complete face or pair of faces each time. This means the present work can only be seen through the lens of Gestaltism as it does not foster any feature-specific potential explanations of face perception. As such, results could be potentially significantly confounded should an individual display a strong preference for a given facial feature such as the nose when providing similarity ratings, as instead of giving an overall similarity estimate between two individual faces, they are instead providing a similarity estimate between two individual's noses in the case of the face-space estimation task. Behaviour such as this, however, could be identified and removed by the deployment of an eye-tracking procedure which would provide evidence for abnormal facial scanning by participants.

Overall, this speaks to a larger overarching point that it is difficult to discern precisely what an individual’s idea of ‘similar’ may be. The tasks simply state, “How similar were these two faces?” or “How similar was the face presented to your own?” and provide no further instructions. This leaves the task of how participants decide how to assess similarity very much open to individual interpretation. There are a vast number of potential ways an individual can determine what they deem to be similar, and many may not even be related to the immediate visual information presented. For example, it is quite possible an individual may have a deeply biased view of how similar an individual looks if they happen to look like a close relative they have. Additionally, should this be the case, research suggests should the individual look similar enough, they will even process the stimuli in question differently leading to a significantly biased similarity estimate between the familiar and any unfamiliar face it may be paired with as it would become a theoretical comparison between internal features (familiar) and external features (unfamiliar) (Veres-Injac & Persike, 2009). Although further along in the reproduction pipeline measures are taken to attempt to limit this potential confound and ensure only dimensions containing visual information remain in the reconstructed image, it continues to remain an unfortunate confound within the current experimental method. This behaviour could be identified, however, reasonably simplistically by comparing participants' ratings of facial similarity with the similarity metric the OpenFace algorithm (Amos et al., 2016) provides. As OpenFace (Amos et al., 2016) is reliant upon spatial distances between landmark points as opposed to pre-existing ideas of what similarity may mean at the individual level, comparing these two distinct datasets would serve as a useful tool to identify and remove participants whose assessment of similarity appears to depend largely upon non-visual dimensions such as the example above demonstrates.

Finally, regarding task-based limitations, it is notable that participants are only provided with the numbers 1 – 7 to provide similarity judgments. This is self-evidently not a particularly granular scale with only whole numbers available for selection meaning information may be immediately lost as participants are likely forced into some form of mental rounding when providing judgments. Additionally, much like the issues presented above, it is incredibly difficult to estimate with any true degree of accuracy what a given individual may or may not decide a '7' may or may not pertain to. Presumably, it must be relative to something in the participant's mind, but this is unlikely to be a concrete unchanging rule but rather a fluid concept with research suggesting that individuals' preferences generally appear to shift frequently depending on myriad factors often far beyond the experimenter's control. Equally, it is possible that an individual's idea of what a given number may relate to may deviate significantly within the task itself, as, as the task progresses, and the participant gathers more information a form of mental averaging may begin to present itself. A set of faces presented during the first 200 trials which may have been awarded a rating of '2' could end up being awarded a '4' should they be presented again in the last 200 trails as that sits closer to the median of a 1 – 7 scale and the participant will have a much stronger concept of the full range of similarities available having performed most of the trials.

One significant issue with the current model in terms of future applications is simply the sheer amount of compute time it requires as the model was initially designed to run on a supercomputer. Unfortunately, in its current form, even following the laborious time-consuming tasks, the output of this approach is a long way from instantaneous. Significant additional expenditure and development are required to achieve this goal. For example, the typical number of permutations required is 10,000 in line with (Chang et al., 2017) earlier work, however, at the time of data processing supercomputing facilities were unavailable leading to this more limited variant. However, even in its more limited form, reconstructions take many hours to complete.

Additionally, due to the algorithmic nature of the work, it remains the case that there are severe limitations to assuming the brain itself functions simply as an intricate piece of software. This means it is difficult to conclude that this model can serve as much more than a potential approximation of how faces are represented in memory as opposed to concrete evidence of how faces are ultimately represented in memory. Although it is accepted here that this criticism can indeed be levelled at the entirety of cognitive psychology, at the time of writing this, it remains valid. Currently, all validation approaches to the model have a strong mathematical focus; this means they focus solely on machine-quantifiable information, namely the spatial relationships between images (see ‘shape’ analyses) and colour comparisons (see ‘texture’ analyses). There is a distinct lack of psychological validation regarding the reconstructed images. As such, we do not know if individuals themselves respond to the reconstructions this model produces in any sort of comparable way to how they respond to their own faces. Without this critical information, it is difficult to determine the overall validity of the model as an accurate portrayal of the participant's self-face representation, as without it, it would stand to reason that any form of potential application relating to use with humans would be limited. However, should the features of the self-face reconstructions reliably correlate with other psychologically valid measures, such as individual differences in psychopathology, for example, the tool may still have a potential use-case.

One potential avenue of further exploration could be within the clinical domain such as the models' potential applications in relation to Body Dysmorphic disorder. The treatment for Body Dysmorphic Disorder (BDD) currently typically involves a combination of psychotherapy, medication, and support. A modified form of Cognitive Behavioural Therapy, known as Cognitive Behavioural Therapy for Body Dysmorphic Disorder (CBT-BDD), is currently considered the most effective psychotherapy for BDD (J.C. et al., 1995). CBT-BDD primarily targets the distorted thoughts and beliefs that individuals with BDD have about their appearance. It helps them recognize and challenge these negative perceptions. Ultimately, CBT-BDD aims to provide individuals with tools and strategies they can use throughout their lives to manage their BDD symptoms effectively. It can lead to long-term improvements in self-esteem, reduced compulsive behaviours, and an enhanced quality of life. Overall, CBT-BDD is a structured and evidence-based approach specifically designed to address the unique challenges faced by individuals with Body Dysmorphic Disorder (Fang et al., 2020).

The efficiency or effectiveness of CBT-BDD is typically assessed using various methods and outcome measures such as symptom severity scales (Yale-Brown Obsessive Compulsive Scale Modified for Body Dysmorphic Disorder), clinical interviews, self-report questionnaires, long-term follow-ups, and general client feedback. Notably, however, none of these assessment measures demonstrate any form of perceptual component and instead largely focus on individuals' subjective beliefs and emotions towards their appearance. It is argued that the evidence starkly suggests that BDD has a distinct perceptual component (Beilharz et al., 2017) as a disorder but any change to this aspect of the presentation of the disorder is not assessed via conventional therapeutic assessment techniques. As such, it is possible patients in some cases may simply be learning to verbalize more realistic beliefs about their appearance but still hold severely distorted images of themselves in visual long-term memory which could play a role in potential relapses. (Harrison et al., 2016).

The model presented in this thesis could potentially serve as a unique measure of perceptual self-representation distortions in BDD or related disorders. Specifically, its inclusion alongside the more traditional self-report measures would form a multidimensional assessment battery, ensuring cognitive, affective, and perceptual body representation distortion in patients were assessed equally. This could ultimately inform more effective treatment strategies which target both the perceptual and conceptual/affective components of the self in this disorder, providing enough evidence emergences supporting the claim that its output is an objective measure of the self-face representation. Furthermore, as work by (Maister et al., 2021) has suggested that, using the reverse correlation method (an analogue of this work), individuals exhibiting higher social self-esteem produce more accurate 'self-portraits', with the 'self-portraits' produced by the method arguably being an adequate reflection of the participant's self-face representation. As such, models such as these may provide a valuable insight into the patients' inner state and even serve a potential predictive tool to assess the likelihood of individuals going on to develop more severe symptoms with self-esteem also serving as a predictor of both onset and severity of Anorexia Nervosa, Bulimia Nervosa, Social anxiety disorders and depressive disorders.

Conclusion:

In conclusion, the project aimed to develop a novel method for visualizing the self-face representation as stored in visual long-term memory. The method itself was successfully tested using ideal theoretical observers based upon a face detection algorithm; however, unfortunately, results from a small pilot on human subjects were not significant. Notably, however, the statistical tests deployed on human data were underpowered, with generic power sample estimates sitting at 23 subjects for a standard paired samples t-test to achieve a potential effect size of 0.8. The tests utilized in this study (involving human participants), however, only included 8 subjects. Additionally, some analysis performed however (Welch's t-test) did include a larger sample ($N_1$=18, $N_2$=60, $\sigma_1$=0.45, $\sigma_2$=0.40) which did return a significant result and was adequately powered to detect $\delta$=0.40 between population means over 10^6 simulations (0.90, roughly 90%). One alternative approach to face-space estimation was also tested, but results remained largely inconclusive.

The Face space, which is modelled here, is a theoretical framework used to represent how humans perceive and recognize faces. It posits that faces are processed within a multidimensional psychological space, where faces are located based on their features and characteristics (Valentine et al., 2016).

At the core of this model sits a loosely modified version of "norm-based coding" theory. This theory suggests that faces are encoded relative to a prototype or average face, with deviations from this prototype determining individual face recognition.

The stability of face space over time, however, is the subject of ongoing research and debate in the field of cognitive psychology. There is evidence to suggest that face representations within face space can be relatively stable but can also be influenced by various factors (Rhodes, 2017).

Factors that may influence the stability of the face space over time may include experience-dependent changes. For example, some individuals may be markedly better at recognizing faces belonging to a certain ethnic group due to exposure during early years. This is a phenomenon best exemplified by the ‘other race’ effect (Li et al., 2015).

Other factors could include adaptive changes as the brain adapts to new perceptual challenges, potentially shifting its focus onto different facial features should certain aspects remain covered, for example, in predominantly Muslim countries where hijabs are the norm or after the massively widespread use of face masks during the COVID-19 pandemic. General individual differences, such as the variability of face recognition abilities in general, genetic components, and personal experiences, also likely play a role. Despite this, however, there is evidence to suggest that certain aspects of face space representations remain relatively stable over time, particularly core facial identity features that are essential for recognizing individuals (Abudarham & Yovel, 2016).

Overall, although the evidence remains somewhat mixed, it would appear that at least some aspects of the face space appear to be quite a malleable construct. As such, it is likely this potential instability could greatly affect the use of the current method as it does largely depend upon the face space itself, at least at the level of the small portion being modelled being stable and robust. This could perhaps help to explain some of the discrepancy between findings and go some way to explaining the seemingly poor test-retest results exhibited here. Furthermore, should the model remain at least largely theoretically correct, this would strongly suggest that the self-face representation itself that individuals hold is a malleable construct and not fixed as a snapshot in time like a photograph is. It would stand to reason that as both external and internal factors affect the arrangement and constitution of the face space on a moment-by-moment basis, the representation of the self-face also likely shifts at the point of view of the observer. Notably, however, it is highly unlikely this shift is ever severe enough that it becomes utterly unrecognizable to the individual with evidence like that cited above suggesting that despite this potential rearrangement core facial identity processes will remain stable and unchanged.

One method to accelerate the process of obtaining similarity estimates which appear to be a major pitfall of the initial approach is via ‘Spatial Arrangement method’ or SPaM as demonstrated earlier in this work, which allows participants to theoretically provide multiple judgments in each trial by taking spatial measurements between stimuli which participants are asked to arrange within a given area onscreen (Hout et al., 2013).

Although it has been suggested that there are caveats to this method such as it essentially disregarding feature representations and ultimately yielding slightly less reliable data than the pairwise method which is extensively reviewed elsewhere in this document. It does produce results that are broadly comparable on average when both methods are compared according to some papers (Verheyen et al., 2022) and is, of course, significantly faster. Notably however these results where not replicated here despite a sample size of (N=28 (F), N=39(total) being obtained. This could be partially explained perhaps by $S_E$ = 0.16 within the largest (N=39) sample however.

Regarding the current project, it is also notable that the original pairwise method did not, in fact, stand up to test-retest scrutiny based on results of Experiment 2 with non-significant correlations shown within participant ratings when asked to perform the same task multiple times over several days. The 'SPaM' method also failed to show an adequate degree of correlation between tasks. As such it is difficult to conclude from these results in isolation that the paradigm(s) deployed here are mathematically comparable although sample sizes where limited throughout this project in no small part due to the use of opportunity sampling. Furthermore, results from 'Experient 1' largely failed to demonstrate a significant difference between means and presented myriad methodological concerns discussed elsewhere in this work, however from a purely mathematical perspective they too remain largely underpowered.

Although results from the new self-face representation reconstruction task have been somewhat mixed for human participants, the task does open several potential avenues for interesting further research. Additional lines of inquiry could focus on psychological validation which would aim to demonstrate if participants detected any psychological similarity between themselves and their reconstructions via a reaction time paradigm such as (Bortolon & Raffard, 2018), and Neurophysiological validation of model output via EEG signal analysis (Gwinn et al., 2021; Knyazev, 2013; Nemrodov et al., 2018). Furthermore, an in-depth exploration of individual differences (providing validation is adequately met) may perhaps prove fruitful.

**Appendix:**

Please contact Dr. Lara Maister or Michael Vinsome* to review the computational work and all additional materials associated with this project.

**References:**

Abudarham, N., & Yovel, G. (2016). *Reverse engineering the face space: Discovering the critical features for face identification.* Journal of Vision, 16(3). https://doi.org/10.1167/16.3.40

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

Anonim. (2019). *dlib C++ Library.* In 10 March 2019.

Ballico, E., Favacchio, G., Guardo, E., & Milazzo, L. (2021). *Steiner systems and configurations of points.* Designs, Codes, and Cryptography, 89(2). https://doi.org/10.1007/s10623-020-00815-x

Beilharz, F., Castle, D. J., Grace, S., & Rossell, S. L. (2017). *A systematic review of visual processing and associated treatments in body dysmorphic disorder.* In Acta Psychiatrica Scandinavica (Vol. 136, Issue 1). https://doi.org/10.1111/acps.12705

Bortolon, C., & Raffard, S. (2018). *Self-face advantage over familiar and unfamiliar faces: A three-level meta-analytic approach.* In Psychonomic Bulletin and Review (Vol. 25, Issue 4). https://doi.org/10.3758/s13423-018-1487-9

Botvinick, M., & Cohen, J. (1998). *Rubber hands “feel” touch that eyes see*

*[8].* In Nature (Vol. 391, Issue 6669). https://doi.org/10.1038/35784

Bushnell, I. W. R. (2001). *Mother's Face Recognition in Newborn Infants: Learning and Memory.* Infant and Child Development, 10(1–2). https://doi.org/10.1002/icd.248

Caharel, S., Poiroux, S., Bernard, C., Thibaut, F., Lalonde, R., & Rebai, M. (2002). *ERPs associated with familiarity and degree of familiarity during face recognition.* International Journal of Neuroscience, 112(12). https://doi.org/10.1080/00207450290158368

Campos, A. (2011). *Internal consistency and construct validity of two versions of the revised vividness of Visual Imagery Questionnaire.* Perceptual and Motor Skills, 113(2). https://doi.org/10.2466/04.22.PMS.113.5.454-460

Chang, C. H., Nemrodov, D., Lee, A. C. H., & Nestor, A. (2017). *Memory and Perception-based Facial Image Reconstruction.* Scientific Reports, 7(1). https://doi.org/10.1038/s41598-017-06585-2

Dotsch, R., & Todorov, A. (2012). *Reverse Correlating Social Face Perception.* Social Psychological and Personality Science, 3(5). https://doi.org/10.1177/1948550611430272

(2021). *Faces as Gestalt Stimuli: Process Characteristics.* In Computational, Geometric, and Process Perspectives on Facial Cognition. https://doi.org/10.4324/9781410612328-12

Fang, A., Steketee, G., Keshaviah, A., Didie, E., Phillips, K. A., & Wilhelm, S. (2020). *Mechanisms of Change in Cognitive Behavioural Therapy for Body Dysmorphic Disorder.* Cognitive Therapy and Research, 44(3). https://doi.org/10.1007/s10608-020-10080-w

Felisberti, F. M., & Musholt, K. (2014). *Self-face perception: Individual differences and discrepancies associated with mental self-face representation, attractiveness, and self-esteem.* Psychology and Neuroscience, 7(2). https://doi.org/10.3922/j.psns.2014.013

Fine, S., Haley, G., Gilbert, M., & Forth, A. (1993). *Self-image as a Predictor of Outcome in Adolescent Major Depressive Disorder.* Journal of Child Psychology and Psychiatry, 34(8). https://doi.org/10.1111/j.1469-7610.1993.tb02098.x

Golaszewski, S., Frey, V., Thomschewski, A., Sebastianelli, L., Versace, V., Saltuari, L., Trinka, E., & Nardone, R. (2021). *Neural mechanisms underlying the Rubber Hand Illusion: A systematic review of related neurophysiological studies.* In Brain and Behaviour (Vol. 11, Issue 8). https://doi.org/10.1002/brb3.2124

Harrison, A., Fernández de la Cruz, L., Enander, J., Radua, J., & Mataix-Cols, D. (2016). *Cognitive-behavioural therapy for body dysmorphic disorder: A systematic review and meta-analysis of randomized controlled trials.* In

Clinical Psychology Review (Vol. 48). https://doi.org/10.1016/j.cpr.2016.05.007

Henry, G. K. (2023). *The Continuous Visual Memory Test: Update and extension on the operating characteristics as an embedded measure of cognitive performance validity.* Clinical Neuropsychologist, 37(1). https://doi.org/10.1080/13854046.2021.2010807

Hirot, F., Lesage, M., Pedron, L., Meyer, I., Thomas, P., Cottencin, O., & Guardia, D. (2016). *Impaired processing of self-face recognition in anorexia nervosa.* Eating and Weight Disorders, 21(1). https://doi.org/10.1007/s40519-015-0223-y

Hopwood, C. J., Bagby, R. M., Gralnick, T., Ro, E., Ruggero, C., Mullins-Sweatt, S., Kotov, R., Bach, B., Cicero, D. C., Krueger, R. F., Patrick, C. J., Chmielewski, M., DeYoung, C. G., Docherty, A. R., Eaton, N. R., Forbush, K. T., Ivanova, M. Y., Latzman, R. D., Pincus, A. L., … Zimmermann, J. (2020). *Integrating psychotherapy with the hierarchical taxonomy of psychopathology (HiTOP).* Journal of Psychotherapy Integration, 30(4). https://doi.org/10.1037/int0000156

Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). *The versatility of SpAM: A fast, efficient, spatial method of data collection for multidimensional scaling.* Journal of Experimental Psychology: General, 142(1).

https://doi.org/10.1037/a0028860

Itulua-Abumere, F. (2013). *Anorexia Nervosa and Bulimia Nervosa Critical Analysis of It's Treatment: Implications and Interventions.* The Open Access Journal of Science and Technology, 1. https://doi.org/10.11131/2013/100007

Jackson, M. C., & Raymond, J. E. (2008). *Familiarity Enhances Visual Working Memory for Faces.* Journal of Experimental Psychology: Human Perception and Performance, 34(3). https://doi.org/10.1037/0096-1523.34.3.556

J.C., R., J., R., & P., O. (1995). *Cognitive-behavioural body image therapy for body dysmorphic disorder.* Journal of Consulting and Clinical Psychology, 63(2).

Kanavakis, G., Halazonetis, D., Katsaros, C., & Gkantidis, N. (2021). *Facial shape affects self-perceived facial attractiveness.* PLoS ONE, 16(2 February). https://doi.org/10.1371/journal.pone.0245557

Karremans, J. C., Dotsch, R., & Corneille, O. (2011). *Romantic relationship status biases memory of faces of attractive opposite-sex others: Evidence from a reverse-correlation paradigm.* Cognition, 121(3). https://doi.org/10.1016/j.cognition.2011.07.008

Knyazev, G. G. (2013). *EEG correlates of self-referential processing.* In Frontiers in Human Neuroscience (Issue MAY).

https://doi.org/10.3389/fnhum.2013.00264

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). *Prior expectations bias sensory representations in visual cortex.* Journal of Neuroscience, 33(41). https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kramer, R., Duran, K., Soder, H., Applegate, L., Youssef, A., Criscione, M., & Keenan, J. P. (2020). *The Special Brain: Subclinical Grandiose Narcissism and Self-Face Recognition in the Right Prefrontal Cortex.* American Journal of Psychology, 133(4). https://doi.org/10.5406/AMERJPSYC.133.4.0487

Kramer, R. S. S., Jenkins, R., & Burton, A. M. (2017). *InterFace: A software package for face image warping, averaging, and principal components analysis.* Behaviour Research Methods, 49(6). https://doi.org/10.3758/s13428-016-0837-7

Kriegeskorte, N., & Mur, M. (2012). *Inverse MDS: Inferring dissimilarity structure from multiple item arrangements.* Frontiers in Psychology, 3(JUL). https://doi.org/10.3389/fpsyg.2012.00245

Lee, H., Ahn, J. K., & Kwon, J. H. (2019). *Effects of Self-Image on Anxiety, Judgement Bias, and Emotion Regulation in Social Anxiety Disorder.* Behavioural and Cognitive Psychotherapy, 47(1). https://doi.org/10.1017/S135246581800022X

Lee, J. K. W., Janssen, S. M. J., & Estudillo, A. J. (2022). *A featural account* *for own-face processing? Looking for support from face inversion, composite face, and part-whole tasks.* I-Perception, 13(4). https://doi.org/10.1177/20416695221111409

Li, Y., Li, W., & Wang, Z. (2015). *Visual Working Memory Capacity for Own- and Other-race Faces: Effects of Set Size and Face Features.* Journal of Vision, 15(12). https://doi.org/10.1167/15.12.670

Maister, L., De Beukelaer, S., Longo, M. R., & Tsakiris, M. (2021). *The Self in the Mind's Eye: Revealing How We Truly See Ourselves Through Reverse Correlation.* Psychological Science, 32(12). https://doi.org/10.1177/09567976211018618

Michel, M., & Peters, M. A. K. (2021). *Confirmation bias without rhyme or reason.* Synthese, 199(1–2). https://doi.org/10.1007/s11229-020-02910-x

Mirams, L., Poliakoff, E., Zandstra, E. H., Hoeksma, M., Thomas, A., & El-Deredy, W. (2014). *Feeling bad and looking worse: Negative affect is associated with reduced perceptions of face-healthiness.* PLoS ONE, 9(9). https://doi.org/10.1371/journal.pone.0107912

Mitchell, D. (2017). *Body Dysmorphia and the Phenomenology of Embodiment.* Journal of the British Society for Phenomenology, 48(1). https://doi.org/10.1080/00071773.2016.1181366

Moon, K., Kim, S., Kim, J., Kim, H., Ko, Y-G. (2020). *Visualizing Mental Representations of Self Through Reverse Correlation.* Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01149

Nestor, A., Plaut, D. C., & Behrmann, M. (2016). *Feature-based face representations and image reconstruction from behavioural and neural data.* Proceedings of the National Academy of Sciences of the United States of America, 113(2). https://doi.org/10.1073/pnas.1514551112

Ninomiya, H., Onitsuka, T., Chen, C. H., Sato, E., & Tashiro, N. (1998). *P300 in response to the subject's own face.* Psychiatry and Clinical Neurosciences, 52(5). https://doi.org/10.1046/j.1440-1819.1998.00445.x

Oh, D. W., Walker, M., & Freeman, J. B. (2021). *Person knowledge shapes face identity perception.* Cognition, 217. https://doi.org/10.1016/j.cognition.2021.104889

Oliveira, M., Garcia-Marques, T., Dotsch, R., & Garcia-Marques, L. (2019). *Dominance and competence face to face: Dissociations obtained with a reverse correlation approach.* European Journal of Social Psychology, 49(5). https://doi.org/10.1002/ejsp.2569

Parkington, K. B., & Itier, R. J. (2018). *One versus two eyes makes a difference! Early face perception is modulated by featural fixation and feature context.* Cortex, 109. https://doi.org/10.1016/j.cortex.2018.08.025

Petersen, L. A., & Leue, A. (2022). *Face Memory and Face Matching: Internal Consistency and Test-Retest Reliability for the CFMT+ and the GFMT-S.* Journal of Individual Differences, 43(3). https://doi.org/10.1027/1614-0001/a000361

Petkova, V. I., Björnsdotter, M., Gentile, G., Jonsson, T., Li, T. Q., & Ehrsson, H. H. (2011). *From part- to whole-body ownership in the multisensory brain.* Current Biology, 21(13). https://doi.org/10.1016/j.cub.2011.05.022

Poltoratski, S., Kay, K., Finzi, D., & Grill-Spector, K. (2021). *Holistic face recognition is an emergent phenomenon of spatial processing in face-selective regions.* Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-24806-1

Potthoff, J., & Schienle, A. (2021). *Effects of self-esteem on self-viewing: An eye-tracking investigation on mirror gazing.* Behavioral Sciences, 11(12). https://doi.org/10.3390/bs11120164

Prnjak, K., Jukic, I., Mitchison, D., Griffiths, S., & Hay, P. (2022). *Body image as a multidimensional concept: A systematic review of body image facets in eating disorders and muscle dysmorphia.* Body Image, 42. https://doi.org/10.1016/j.bodyim.2022.07.006

Rhodes, G. (2017). *Adaptive Coding and Face Recognition.* Current Directions in Psychological Science, 26(3).

https://doi.org/10.1177/0963721417692786

Roseboom, W., & Lush, P. (2022). *Serious Problems with Interpreting Rubber Hand "Illusion" Experiments.* Collabra: Psychology, 8(1). https://doi.org/10.1525/collabra.32274

Sugiura, M., Sassa, Y., Jeong, H., Wakusawa, K., Horie, K., Sato, S., & Kawashima, R. (2012). *Self-face recognition in social context.* Human Brain Mapping, 33(6). https://doi.org/10.1002/hbm.21290

Sui, J., & Humphreys, G. W. (2015). *The Integrative Self: How Self-Reference Integrates Perception and Memory.* Trends in Cognitive Sciences, 19(12). https://doi.org/10.1016/j.tics.2015.08.015

Tajadura-Jiménez, A., Longo, M. R., Coleman, R., & Tsakiris, M. (2012). *The person in the mirror: Using the enfacement illusion to investigate the experiential structure of self-identification.* Consciousness and Cognition, 21(4). https://doi.org/10.1016/j.concog.2012.10.004

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). *Face-space: A unifying concept in face recognition research.* Quarterly Journal of Experimental Psychology, 69(10). https://doi.org/10.1080/17470218.2014.990392

Veres-Injac, B., & Persike, M. (2009). *Recognition of briefly presented familiar and unfamiliar faces.* Psihologija, 42(1), 47–66. https://doi.org/10.2298/psi0901047v

Verheyen, S., White, A., & Storms, G. (2022). *A Comparison of the Spatial Arrangement Method and the Total-Set Pairwise Rating Method for Obtaining Similarity Data in the Conceptual Domain.* Multivariate Behavioral Research, 57(2–3). https://doi.org/10.1080/00273171.2020.1857216

Wiese, H., Hobden, G., Siilbek, E., Martignac, V., Flack, T. R., Ritchie, K. L., Young, A. W., & Burton, A. M. (2021). *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. https://doi.org/10.1037/xlm0001063

Wirawani Kamarulzaman, & Prof Dr. Mohamad Sahari Nordin. (2012). *Confirmatory Factor Analysis On The Big 5 Personality Test Inventory.* 2nd Southeast Asia Psychology Conference.