CellPress REVIEWS

Review

The Face of Image Reconstruction: Progress, Pitfalls, Prospects

Adrian Nestor,[1,*] Andy C.H. Lee,[1,2] David C. Plaut,[3,4] and Marlene Behrmann[3,4]

Recent research has demonstrated that neural and behavioral data acquired in response to viewing face images can be used to reconstruct the images themselves. However, the theoretical implications, promises, and challenges of this direction of research remain unclear. We evaluate the potential of this research for elucidating the visual representations underlying face recognition. Specifically, we outline complementary and converging accounts of the visual content, the representational structure, and the neural dynamics of face processing. We illustrate how this research addresses fundamental questions in the study of normal and impaired face recognition, and how image reconstruction provides a powerful framework for uncovering face representations, for unifying multiple types of empirical data, and for facilitating both theoretical and methodological progress.

Highlights

The visual content of human face representations can be derived through image reconstruction at an unprecedented level of detail.

The neural underpinnings of psychological face representations can be elucidated via image reconstruction from behavioral and multiple types of neural data.

Facial image reconstruction capitalizes on the structure of face space and also serves to validate its representational properties.

The recovery of face representations from both perception and memory can help to clarify the relationship between these two components of cognition.

Image reconstruction can pinpoint the loss of information and/or misrepresentations in face perception across a broad population, including individuals with deficits in face processing.

From Neural Decoding to Image Reconstruction

A core endeavor in the cognitive neuroscience of face perception is to elucidate the neural and psychological representations that subserve face recognition [1–4]. Accordingly, neural decoding has been extensively used to investigate face perception by relating the structure of neural patterns to the properties of viewed faces. For instance, facial identity decoding has shed light on the cortical locus [5–9], the temporal dynamics [10–13] and the broad characteristics of face representations, such as the relationship between identity and expression information [14]. However, the finer-grained content and structure of face representations remain to be uncovered. Further, the ability of such representations to support multiple aspects of cognition, including perception, imagery, and memory, remains to be clarified.

An increasingly popular type of decoding approach has shown promise in addressing the challenges above. Specifically, image reconstruction encompasses a newer set of methods (Box 1) that convert empirical data associated with visual processing into images that capture the visual properties of underlying representations. As a general strategy, reconstruction relies on combining a set of visual features, weighted by the amplitude of specific neural and/or behavioral responses, into images. Beyond this general strategy, however, reconstruction methods vary widely in the types of features utilized, corresponding to different mechanistic hypotheses [15,16], as well as in the procedure for feature combination and in their mapping onto empirical data.

To date, reconstruction endeavors have relied primarily on fMRI to retrieve the content of natural images [17–20], movies [21,22], and of specific stimulus categories such as orthographic characters [23–26] and scenes [27]. In the study of face perception, proof of concept has been provided for the recovery of face images [28] and identifiable facial appearance [29] from fMRI patterns. More recently, facial image reconstruction has been conducted with other data types, including behavioral [30,31] and electroencephalography (EEG) data in humans [32,33] as well as single-unit recordings in monkeys [34]. Further approaches based on neural networks have

[1]Department of Psychology at Scarborough, University of Toronto, Toronto, Ontario, Canada
[2]Rotman Research Institute, Baycrest Centre, Toronto, Ontario, Canada
[3]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA
[4]Carnegie Mellon Neuroscience Institute, Pittsburgh, PA, USA

*Correspondence
adrian.nestor@utoronto.ca (A. Nestor).

Box 1. Neural Decoding, Image Reconstruction, and Result Validation

Image reconstruction illustrates an increasingly popular type of decoding used in the analysis of neural data. In the study of perception, decoding serves to predict stimulus properties from different types of neural recordings. For instance, linear classifiers have been extensively used to decode facial identity, expression, and orientation from fMRI data associated with viewing faces [108]. Reconstruction evinces a more ambitious aim of recovering not only broad stimulus properties but also an approximation of the stimulus itself by predicting its detailed perceptual properties, such as precise shape and color parameters. In turn, broader stimulus properties may emerge from reconstruction outcomes – for instance, gender and identity are perceptual properties that are visualizable in a given reconstruction [36]. Similarly to other decoding techniques, reconstruction operates on neural patterns and is well suited to capturing distributed representations. However, reconstruction tends to be more computationally intensive because it targets a larger number of variables for prediction purposes (e.g., the values of different visual features and, ultimately, the intensity value of every pixel in an image).

Thus, unlike other forms of decoding, reconstruction is, by design, more thorough and fine-grained. Importantly, however, its aim is not necessarily to recover a perfect replica of a stimulus [109]. If the purpose of visual representations is ultimately to facilitate appropriate behavioral responses [110], rather than to thoroughly capture the visual world, then reconstruction cannot generally be expected to fully recover a stimulus. Hence, the comparison between reconstructions and corresponding stimuli provides, when available, an immediate but incomplete means of evaluating reconstruction results. Additional validation tools, such as behavior-based measures of reconstruction accuracy by human observers, can be relatively valuable in this sense and are commonly used for assessing facial image reconstruction [31,33,36]. Further, regarding memory-based reconstructions, asking participants to recognize their own reconstructions, without appeal to direct comparisons with experimental stimuli, can provide an important and exacting test of reconstruction success [30,78]. Clearly, however, future efforts should be directed to the design not only of better reconstruction algorithms but also of additional, rigorous validation procedures.

Of course, the primary goal of reconstruction may not always be the study of visual representations, and the evaluation of reconstruction outcomes would then need to consider factors beyond representational accuracy – for instance, improving reconstruction quality through dedicated models that boost usability and utility would be highly relevant in the design of reconstruction-based brain–computer interfaces.

aimed to improve the accuracy and image quality of facial stimulus reconstruction [35,36]. This new body of work can benefit from a close examination to uncover not only its methodological progress but also its theoretical implications, its merits, and its challenges.

The present review focuses on faces, as a target domain, to ensure even ground for the comparison of a variety of reconstruction approaches. Notably, faces provide a restricted but important testbed because they provide the epitome of visual expertise while also exhibiting a well-constrained visual structure (e.g., by sharing the same parts in similar configurations) [37]. Historically, their use as a target domain has played a seminal role in the development and validation of decoding techniques, such as those based on multivoxel pattern analysis [5,38]. Similarly, image reconstruction could benefit from a targeted assessment in the context of this visual domain.

In the current review, we show how recent investigations provide a coherent and informative perspective on face processing. Notably, these studies open up new avenues for researching memory representations and visual distortions associated with specific deficits. At the same time, we consider the need to distinguish between recovering the appearance of a stimulus and uncovering the genuine content of visual representations. Accordingly, we attempt to clarify and illustrate how image reconstruction does not reflect only the properties of face stimuli but, crucially, also those of neural and psychological representations.

From Image Reconstruction to Perceptual Representations

The feasibility of facial image reconstruction has served as a powerful demonstration of the richness of visual content underlying face perception and of the ability to extract this content from neural signals [28]. Because face identification stands out as the hallmark of face recognition, reconstruction has the opportunity to shed new light on representations of facial identity. We consider here three goals that are essential for rendering reconstruction relevant to the study of such

Glossary

Action units (AUs): facial movements, each related to a particular muscle or group of muscles, as described by the facial action coding system (FACS). AUs have proved to be particularly useful in the physical description and recognition of facial expressions. For example, AU 6 (cheek raiser) is often used to express sadness whereas AU 25 (lips part) is used to express joy.

Deep convolutional neural networks (DCNNs): artificial networks consisting of multiple layers of simulated neurons. When applied to image processing, DCNNs begin with raw images and recode information across layers by convolution (i.e., a linear combination of neighboring pixels according to the weights expressed in a filter). Convolution implements filters with a restricted spatial receptive field, inspired by neural processing in the visual cortex. The hierarchical nature of DCNNs is also broadly inspired by the organization of the visual cortex.

Eigenfaces: components derived from the application of principal component analysis (PCA) to a set of images. PCA can reduce the dimensionality of a dataset by encoding it via a relatively small number of components that capture the main forms of variability in the data. Conversely, a face image can be recovered fairly well from a weighted combination of a limited number of eigenfaces.

Ensemble coding: a type of visual coding that is deployed for groups of similar objects in the visual environment: the leaves of a tree, a basket of apples, a crowd of people. To handle such a wealth of information, the visual system is believed to capitalize on its redundancy. Specifically, it may encode a statistical summary representation of all the elements, an average apple or an average face, instead of every single apple or face.

Invariant face representations: types of representations that encode visual information independently, to some degree, of variability in individual appearance. Considerable variability in the projection of a face image on the retina occurs owing to extrinsic factors such as angle of view and lighting, as well as to intrinsic factors such as expression and aging.

Mental imagery: visual experience whose content does not directly derive from an external stimulus, as takes place representations: (i) relating reconstructions to the architecture of a representational space and to behavioral performance; (ii) deriving/analyzing reconstructions into fundamental components of face representations, such as shape and surface properties, and (iii) examining the robustness of reconstruction over image variability (e.g., invariance to viewpoint and expression).

The concept of face space (Box 2), a multidimensional space within which the distance between faces corresponds to their perceived dissimilarity [39], is fundamental to our understanding of face representations [40–43]. Accordingly, recent work [29] has not only related reconstructions directly to face space but has grounded them in its architecture based on the structure of fMRI patterns and behavioral data (i.e., ratings of pairwise face similarity). In this work, image reconstruction was achieved in a series of steps (Figure 1): first, by assessing the representational similarity [44,45] across faces from neural and behavioral data; second, by converting similarity estimates into face space constructs; third, by synthesizing visual features from the structure of this space; and, last, by linearly combining such features into reconstructions of target stimuli. The twofold success of deriving identifiable face information from fMRI patterns and from behavioral data demonstrated the ability of reconstruction to uncover the visual codes structuring the architecture of face space. This demonstration was reinforced by the correspondence between behavioral and neural-based face spaces and, also, between their respective reconstruction outcomes.

Of note, the earlier demonstration was conducted with faces that are highly homogeneous in appearance (e.g., front views of adult Caucasian male faces). Although reconstruction efforts often consider the wider scope of image variability that is introduced, for realistic stimuli, by race, gender, and pose [28,36], face recognition is remarkable in its ability to discriminate even among highly similar individuals. Hence, synthesizing discriminable images from data associated with homogeneous stimuli provides a complementary strategy that takes direct aim at fine-grained representations while also delivering an exacting test of reconstruction.

during perception, but instead is derived from memory. Imagery and visual working memory are deeply related to each other, although their precise overlap remains to be determined.

Prosopagnosia: a deficit in face recognition abilities, and in identity recognition in particular, that cannot be explained through a more general perceptual or cognitive deficit. Prosopagnosia can occur as a consequence of brain damage in premorbidly normal individuals (acquired prosopagnosia) or as a lifelong developmental disorder (congenital/developmental prosopagnosia).

Box 2. Face Space

The concept of face space has provided a seminal framework for the study of face processing. Originally, face space was introduced as a psychological similarity space [39]. Within this space each face is represented by a location, and the distance between faces corresponds to their perceptual (dis)similarity: more distant faces in this space are more dissimilar to each other.

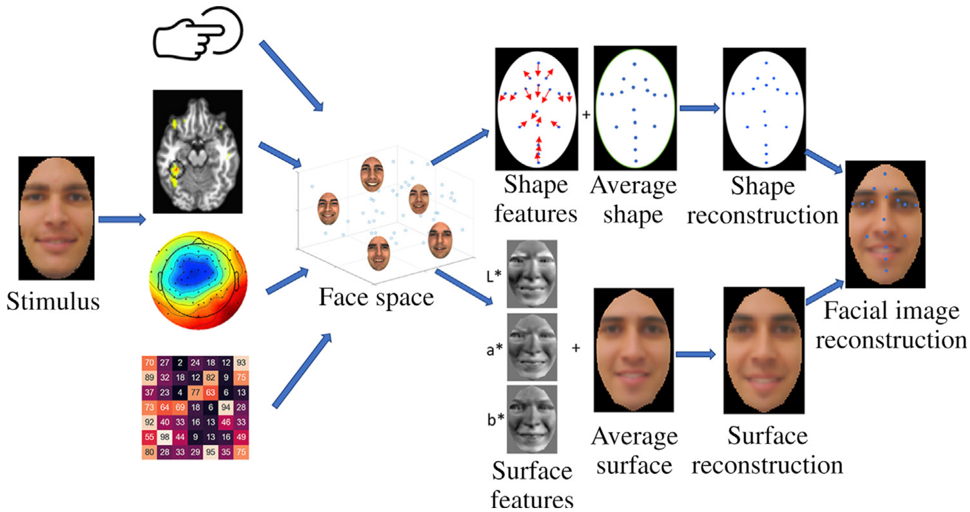
An influential norm-based model postulated that faces are encoded relative to a specific prototypical face (i.e., a norm face) located at the origin of the space, whereas less-typical faces are placed further away from the origin [111]. Alternatively, in an exemplar-based model [112], faces are represented in the space without reference to a central prototype, and the distinctiveness of a face is determined by how densely populated its designated region of space is.

The dimensions of the space correspond to properties or parameters by which faces vary. In the first computational account of face space, dimensions corresponded to eigenfaces derived through the application of PCA to face images [68]. PCA-based models have been useful both in psychology, to explain a plethora of behavioral effects, and in computer vision, as the basis for the first generation of automatic face-recognition systems.

However, an eigenface-based face space is highly limited in its capacity to support and explain face recognition (e.g., because of the inability of its image-based features to deal with invariance). Subsequent advances have been made by applying PCA separately to shape and surface properties rather than directly to images [51]. More recently, a new generation of face-space models have been constructed from latent features encoded on the layers of DCNNs with various architectures and training regimens [3,61]. Modeling neural-based face space across distinct cortical regions with corresponding spaces from distinct DCNN layers is particularly appealing as a way to capitalize on the hierarchical nature of visual processing in both biological and artificial systems.

Overall, computational approaches in the study of face space are crucial for assessing its properties: its dimensionality, its similarity metric, its flexibility to learning, and the nature of its dimensions. Image reconstruction provides a new and promising direction of research in the study of these properties.

CellPress
REVIEWS



Trends in Cognitive Sciences

Figure 1. Facial Image Reconstruction Informs Fundamental Aspects of Face Representation. Notably, reconstruction speaks to the topography of face space and to the separate encoding of shape and surface information. Reconstruction can rely on multiple types of neural and behavioral data as well as directly on image information (e.g., processed by a model or a theoretical observer). An illustration of image reconstruction methodology (adapted, with permission, from [33]) involves estimating a face-space construct from the structure of experimental data, followed by the synthesis of shape and surface features from face space and the combination of such features into image reconstructions of a viewed face (only a representative subset of facial landmarks are displayed for shape; surface features are displayed for L*, a*, and b*, which correspond to the lightness, red-green, and yellow-blue channels of color vision).

An important question, however, is how reconstruction and the underlying face space reflect fundamental properties of face representations. Notably, the segregation of shape and surface information defines a key aspect of visual processing [46], and both types of information support crucial aspects of face recognition [37,47–50]. A landmark study relying on single-unit recordings in macaque brain [34] examined this segregation and its contribution to reconstruction. The results showed that shape and surface features are predominantly encoded by groups of neurons in different patches of inferotemporal (IT) cortex (Figure 2A). An examination of shape and surface features, based on active appearance models [51], demonstrated that such features can be efficiently decoded from neural recordings and integrated into image reconstructions (Figure 2B). Importantly, this study demonstrated that IT neurons exhibit tuning to specific features rather than encoding information about face exemplars (e.g., 'concept cells' in the human hippocampus [52,53]). Theoretically, these neurons function as the axes of a space that flexibly supports the distributed representation of any number of faces (i.e., as linear projections). Conversely, a linear combination of feature/axis values, through the readout of corresponding neurons, enables image reconstruction. Hence, these results establish the neural validity of an axis model of face representation relying on separate sets of shape and surface features.

Nevertheless, a crucial issue for reconstruction concerns the retrieval of **invariant face representations** (see Glossary). Our ability to recognize facial identity despite considerable variation in appearance [4], such as that due to changes in viewpoint and expression, suggests that some aspects of representations and the underlying face space topography are invariant to such changes. Hence, one should assess whether reconstruction captures features that are robust to identity-preserving transformations. Namely, does reconstruction retrieve only lower-level image properties (e.g., average lightness and image contrast) or higher-level facial identity properties? An examination of reconstruction results with front-view and profile faces from single-unit recordings [34] showed that a shared face space topography underlies both sets of

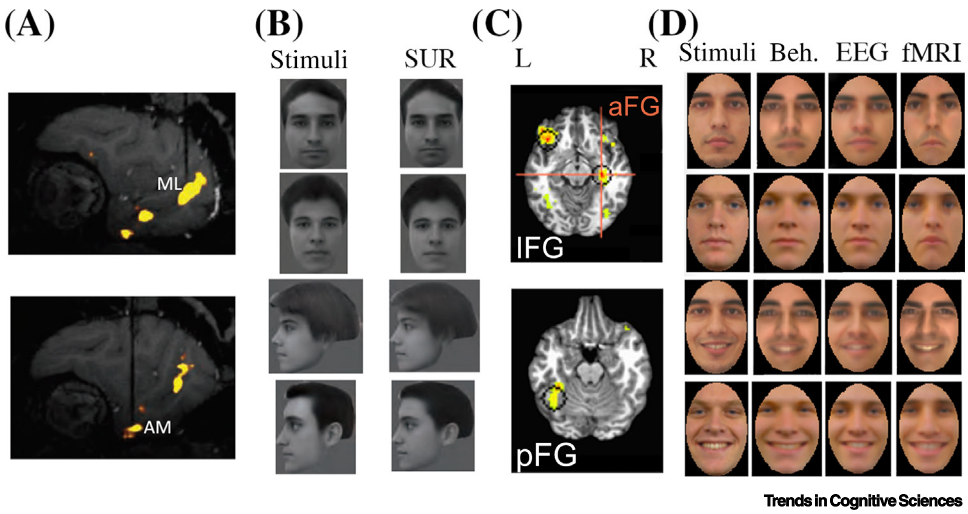


Figure 2. Cortical Areas That Support Identifiable Facial Image Reconstructions across Variation in Viewpoint and Expression. (A) Sagittal slices showing the location of fMRI-mapped face patches in one monkey, and (B) examples of stimuli and their corresponding reconstructions based on single-unit recordings (SURs) in these areas (adapted, with permission, from [34]). (C) Axial slices showing the location of cortical areas in the right anterior fusiform gyrus (aFG) and the left posterior FG (pFG) in humans, and (D) examples of stimuli and their corresponding reconstructions based on multivoxel patterns in these areas – additional reconstructions based on behavioral and electroencephalography (EEG) data are also shown for the same stimuli. Figure adapted, with permission, from [29,33]. Abbreviations: AM, anterior medial face patch; Beh., behavioral reconstruction; IFG, inferior frontal gyrus; ML, middle lateral face patch.

reconstructions. A similar conclusion was reached by an examination of EEG-based reconstructions of faces displaying neutral and happy expressions [32]. These findings strongly point to the ability of reconstruction to capture invariant visual information regarding facial identity.

Thus, image reconstruction concurrently validates and capitalizes on fundamental aspects of face representation: face space topography, the segregation of different visual properties, and the use of invariant features for recognition purposes. Accordingly, reconstruction may be used to probe an entire range of aspects regarding face representations, such as their spatiotemporal profiles and their computational underpinnings.

Image Reconstruction: Where, When, and How

Which cortical areas subserve image reconstruction? Current work demonstrates that fusiform gyrus (FG) regions in the human brain, including a region anterior to the fusiform face area, support reconstruction [29] and the extraction of underlying shape and surface features [33] (Figure 2C). Considering the presumed homology between the cortical systems for face processing in the human and monkey brain [54,55], these findings are in broad agreement with the corresponding ability demonstrated for middle and anterior face patches in monkey IT cortex [34]. More importantly, the location of these regions concurs with the encoding of invariant features in high-level visual cortex [56].

Nevertheless, neither face selectivity nor successful classification of facial identity guarantees the ability of a cortical region to support reconstruction. For instance, neither the face-selective occipital face area (OFA) nor inferior frontal gyrus (IFG) areas, that are capable of identity classification, appear to support identifiable reconstructions [29]. Presumably, this reflects the involvement of the OFA in lower-level face processing and the role of IFG areas in higher-level semantic processing [57]. Crucially, these findings illustrate a notable benefit of reconstruction: explicating if and how selectivity and/or classification results rely on relevant visual information as opposed to higher-level conceptual properties or task-related attributes. However, when one considers a broader range of stimuli, encompassing variation due to race, gender, or pose, reconstruction relies on a larger network of regions including, for instance, occipital areas [28,36]. Thus, when reconstruction aims to recover visual properties beyond those strictly needed for identification, then it benefits from capturing both lower- and higher-level information encoded in the larger network for face processing.

A complementary aspect of image reconstruction concerns its time-course as revealed by methods with suitable temporal resolution. In this sense, EEG-based reconstruction has proved to be not only possible but also informative [32]. For instance, both shape and surface features underlying reconstruction can be recovered over an extended period of time, starting ~150 ms after stimulus onset and gradually declining after 300 ms. Interestingly, a multimodal evaluation showed that EEG-based reconstructions are maximally similar to their behavioral and fMRI counterparts (i.e., shape-based reconstruction based on the posterior FG) in the proximity of the N170 event-related component [33] (Figure 2C for examples of reconstruction across modalities). This is consistent with the role of the N170 component in identity processing [58,59], and it accounts for the nature of visual information during this time-window in the EEG signal. More generally, relating reconstruction efforts across modalities is valuable both for broadening our understanding of face representations and for facilitating the mutual validation of different data types.

However, representational and spatiotemporal aspects of face processing speak only indirectly to its neurocomputational mechanisms. Recently, **deep convolutional neural networks** (DCNNs) have been explored as models for human face processing [60–63], given their ability to match human-level performance in face recognition [64]. In particular, the hierarchical nature of DCNNs provides an appealing opportunity to compare successive levels of visual processing between such models and the visual system [65–67]. In parallel, DCNNs have also been used as tools for image reconstruction (Figure 3). For instance, recent work [35] has demonstrated the utility of higher-level representations encoded by higher layers of DCNNs. Specifically, this work used the output of a 14-layer convolutional encoder as latent features for face encoding and decoding. To this end, fMRI responses were related to latent features by inverting a linear mapping of such features onto brain responses. Then, patterns of latent features were converted into reconstructions by inverting the nonlinear transformation from stimuli to features. DCNN features provided superior results compared with **eigenface** features, as obtained through the application of principal component analysis (PCA) to face images [68], as used by previous reconstruction approaches [28]. Similarly, another study [36] used the output of a 13-layer variational autoencoder (VAE) to mediate between fMRI responses and stimulus properties. Again, the appeal to VAE features facilitated reconstructions that were more accurate than those based on eigenfaces, and with better image quality. Further, latent feature representations corresponding to occipitotemporal fMRI patterns supported gender classification and pointed to the possibility of face decoding from imagery.

Although the earlier mentioned studies focused on boosting reconstruction accuracy and image quality, they also speak to the relevance of face spaces encoded on DCNN layers as models of neural face space. Clearly, a full-blown computational account needs to consider the validity of multiple levels of representations in a DCNN together with the biological plausibility of the transformations performed on these representations [65]. Importantly, however, the aforementioned work demonstrates the benefit of including reconstruction as a tool in assessing the validity of DCNN architectures as face processing models.

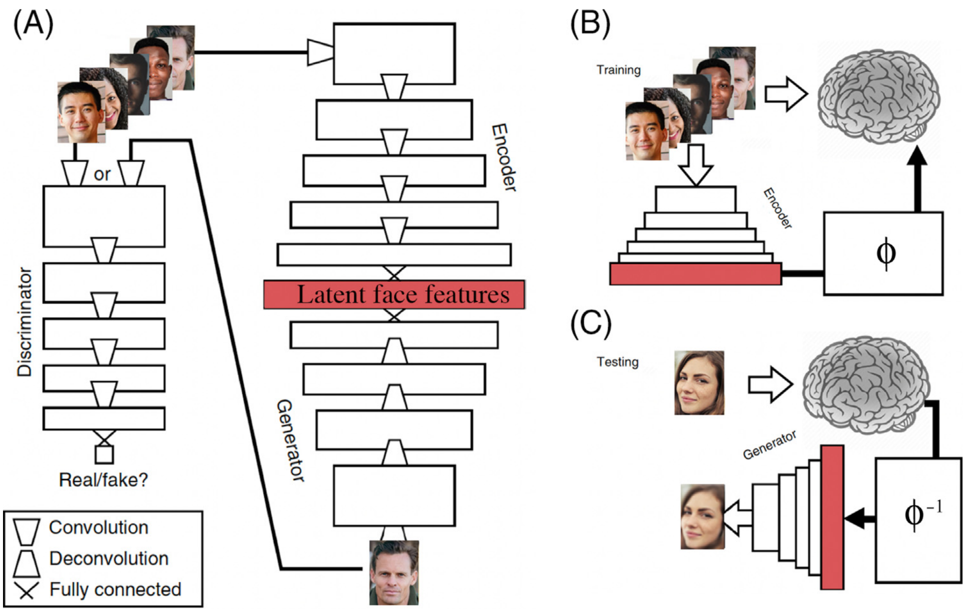


Figure 3. An Illustration of Deep Neural Network Application to Image Reconstruction. (A) An encoder network maps a face image onto latent features, which the generator network converts into a novel face image. The discriminator network evaluates whether a given image, from the original dataset or from the generator output, is real or fake (i.e., a generator output). (B) Encoding associates a latent vector with the corresponding brain response pattern. (C) Encoding is inverted to convert a brain response pattern into an estimate of latent face features which the generator network can translate next into a reconstructed face image. Figure adapted, with permission, from [36].

Perceptual and Memory-Based Face Representations

The relationship between perceptual and mnemonic representations defines a key research area in cognitive neuroscience. For example, extensive work suggests that, during **mental imagery**, visual information is recovered in the ventral pathway via top-down mechanisms for memory retrieval, in contrast to predominantly bottom-up mechanisms employed by perception [69–71]. However, the precise content of memory representations remains to be uncovered. The application of image reconstruction to face memory can be particularly informative given our remarkable ability to remember the appearance of a face. Building on the success of perception-based reconstruction, memory-based reconstruction may elucidate the reliance and the uniqueness of memory representations relative to their perceptual counterparts.

The feasibility of reconstructing face images from working memory has been explored [72] using an eigenface-based approach that linearly maps eigenface features to fMRI patterns (e.g., by fitting a regression model of eigenface scores to multivoxel patterns) [28]. In this study [72], participants imagined the appearance of a target face, singled out by a retro-cue from a pair of face stimuli. Then, corresponding perception and memory-based information was retrieved from the angular gyrus (ANG). Reconstruction outcomes captured information about skin color and emotion, and were thus not limited to low-level visual properties. Importantly, reconstruction was used to refute the hypothesis that the ANG plays only a content-general role in memory retrieval [73] and to confirm that mnemonic content is represented in the ANG during both encoding and retrieval. However, the diversity of the stimuli (e.g., different races, genders etc.) as well as the limited accuracy of memory-based reconstructions precluded definitive conclusions regarding the nature of ANG representations. Namely, although reconstructions were pictorial in format, their success may have reflected the correlation between non-perceptual ANG information and specific eigenfaces (e.g., between semantic information regarding ethnicity and eigenfeatures capturing color variation). In addition, these findings in ANG raise the possibility that other regions, such as medial temporal lobe structures that support face memory and perception [74], may show more promise for such investigations.

Interestingly, behavior-based reconstruction [30] has demonstrated the ability of long-term memory to encode fine-grained visual information. A method relying on similarity ratings between visible stimuli and faces retrieved from long-term memory recovered the appearance of both famous faces and newly learned faces. Stimulus homogeneity and the robustness of the reconstruction results suggested that detailed visual information, rather than broad semantic information, underlies this success. Surprisingly, the accuracy of memory-based reconstruction for unfamiliar faces, learned during the experiment, surpassed that of perception-based reconstruction. Presumably, this indicates that extensive familiarization with a stimulus may lead to memory representations whose richness in diagnostic content is comparable with that of their perceptual counterparts [75].

In a similar vein, by appeal to similarity ratings between stimuli and faces retrieved from memory, another study [31] provided evidence for encoding of 3D shape information in mnemonic representations of personally familiar faces. Further, it showed that face representations did not capture only semantic information regarding gender or age, because such factors were modeled separately in the reconstruction procedure. The robustness of the reconstructions was further evaluated by stringent behavioral tests in which participants were required to recognize reconstructions over changes in viewpoint and age, confirming the ability of the method to extract invariant visual information about identity.

Thus, current work confirms the visual nature of memory face representations, takes steps to uncover its content, and points to its underlying neural resources. Further, it suggests reliance on a representational topography that is shared with perceptual representations [30]. However, the precise relationship between working memory, long-term memory, and perceptual face representations remains to be clarified [72]. In addition, the reliance of these results on restricted samples of healthy participants raises obvious questions regarding their generality.

The Veracity of Face Representations

How veridical are our face representations, and how do they differ across individuals? Image reconstruction provides a unique window into the veracity and idiosyncrasy of face representations both in healthy individuals and in those who suffer from perceptual impairments associated with a variety of visuocognitive deficits.

A particularly relevant impairment is **prosopagnosia**, a deficit in face recognition abilities that is not due to more general problems with vision or memory [76,77]. An investigation of dynamic expression perception in acquired prosopagnosia [78] examined emotion representations in a patient with extensive bilateral occipitotemporal lesions [79]. Consistent with the lesion pattern, face recognition abilities – and identification in particular – were severely impaired. However, regions subserving expression recognition, such as the posterior superior temporal sulcus and the amygdala, were intact, suggesting normal expression representations. To evaluate this hypothesis, expression representations were recovered by relating facial **action units** (AUs) [80] to behavioral performance in an emotion-categorization task. Reconstructions of dynamic expressions were comparable between the patient and age-matched controls, confirming the hypothesis above. Intriguingly, however, the patient was impaired at recognizing static versions of his own reconstructions, suggesting that dynamic and static facial information are subserved by different cortical pathways. These results illustrate a valuable use of reconstruction to characterize expression representations and to elucidate the functional organization of the neural system. They also set the groundwork for future investigations of more thorough representations that include diagnostic surface information, such as color [81], and more complex AU information in compound expressions [82]. However, they do not pinpoint the differences in representational content between the patient and healthy controls that may account for the patient's deficits.

Recent work [83] addressed the challenge of identifying representational differences due to aging. A decline in face recognition abilities associated with healthy aging is well documented [84–86] and is likely to involve degrading representations of identity [87]. To confirm and clarify such representational changes, behavior-based reconstruction was conducted in a group of older adults (aged 62–71 years) and younger controls (aged 18–31 years). The study revealed a loss of representational accuracy in older adults for both shape and surface features that are crucial for recognition, such as eye shape and skin tone [43,88,89]. Interestingly, however, aging accounted only for a small proportion of variance in reconstruction results, pointing to the marked impact of individual differences on face representations.

While the veracity and effectiveness of visual representations come under scrutiny as a window into the nature of visual deficits, non-veridical representations extend beyond the scope of such deficits. An intriguing and pervasive example is that of **ensemble coding**, as documented by extensive behavioral work [90]. Whenever we encounter a group of objects or faces (e.g., a crowd), the visual system seemingly synthesizes a summary representation, by averaging visual properties of individual elements, at the expense of encoding every element. For instance, it may synthesize the representation of an individual with an average expression and identity, although no such individual exists in a crowd [91–93]. To establish the existence of neural summary representations, reconstruction was applied to EEG recordings corresponding to the perception of face ensembles (i.e., groups of faces presented simultaneously) [94]. This investigation recovered summary representations (i.e., the average of individual faces) while showing that information about specific individuals is discarded. It further indicated that ensemble face processing exhibits distinct neural dynamics compared with individual face processing. Thus, reconstruction was instrumental in confirming the neural status of summary representations, in visualizing their appearance, and, more generally, in validating an extensive body of behavioral work.

Taken together, the aforementioned results demonstrate that reconstruction methodology can reveal not only objective stimulus properties, as encoded by our visual system, but also the complexities of visual mental content that may not accurately and entirely reflect the visual world.

Concluding Remarks and Future Directions

The obvious promise of image reconstruction is that of revealing the fine-grained content of visual representations. Nevertheless, its application to face recognition demonstrates considerably more theoretical value and versatility than only a quantitative boost in the detail with which it characterizes representations.

Extending neural-based reconstruction to a behavioral counterpart provides an important opportunity to examine the relationship between brain and behavior [95] in face recognition. Similarly, its applicability to different types of neural data serves to combine their relative strengths and validate each other's conclusions within and across species. Equally important, the use of DCNNs serves a twofold role in modeling visual processing and in facilitating reconstruction.

Outstanding Questions

How do representations of familiar faces differ from those of unfamiliar faces? Extensive work documents differences in recognition performance and in the neural profile of processing familiar versus unfamiliar faces. What aspects of image reconstruction (e.g., content, representational structure, spatiotemporal dynamics) capture these differences, and how do they account for them?

Can we achieve neural-based reconstructions of dynamic facial information with suitable temporal resolution? For instance, can we recover detailed visual representations of dynamic expressions by capitalizing on the high temporal resolution of EEG data?

What does image reconstruction reveal about the other-race effect? Faces of a different race are often recognized more poorly than faces of one own's race. This difficulty presumably reflects the use of less efficient and/or inaccurate visual representations for outgroup faces. What are the specific features impacted in such representations, and how well do they account for the recognition decrements that are characteristic of this effect?

How does perceptual interference alter the structure of mnemonic face representations? For instance, how does a series of face stimuli impact on the representation of a target face recovered from memory? Prior work with simple shapes and colors suggests that visually dissimilar interference erases memory representations, whereas similar interfering items blur these representations. Is face memory equally sensitive or more resilient to such effects?

Can image reconstruction elucidate the heterogeneity of visual deficits in prosopagnosia? Recent work has emphasized the diverse profile of face processing deficits and, presumably, their diverse etiology in congenital prosopagnosia (CP). Can image reconstruction facilitate a better categorization and understanding of different CP (sub)types?

How well does reconstruction methodology targeting faces generalize to other visual categories (e.g., visual words)?

To date, reconstruction encompasses a variety of approaches, as illustrated by their reliance on active appearance features, DCNN latent features, or those synthesized from behavioral and neural data. Nonetheless, despite this variability, current work provides a cohesive perspective on face recognition that emphasizes the reliance on separate shape and surface features, the validity of an axis model of face space, and the encoding of invariant features in high-level visual cortex, as well as a shared representational basis for perception and memory. Further, it demonstrates how changes in visual representations account for decrements in face recognition, and how misrepresentations of objective information help the system to cope with particular tasks (e.g., processing an abundance of redundant information as an aggregate).

Thus, reconstruction has already provided important insights into face recognition. However, its prospects remain at least as impressive. Clearly, the scope of its applicability can be expanded substantially (see Outstanding Questions). A notable challenge is that of characterizing genuine distortions in representations as opposed to only information loss. Misrepresentations of facial appearance in prosopagnosia and prosopometamorphosia [96], negative biases in expression recognition in borderline personality disorder [97] and schizophrenia [98], face illusions and hallucinations [99], as well as perceptual deficits in dementia [100,101], provide rich fields of inquiry that can benefit from this methodology.

Moreover, memory-based reconstruction opens up the possibility of concrete applications, such as neural-based sketch artists. An important challenge in this sense is the amount of data that is necessary for reconstruction purposes for any given individual. Exploiting the mapping between neural recordings and image properties across a larger and more diverse sample of individuals could improve reconstruction via transfer learning [102]. However, such

Box 3. Methodological Paradigms for the Study of Visual Representations

Image reconstruction is related to several methodological paradigms in the study of visual representations. In particular, representational similarity analysis (RSA) has been used extensively to uncover the structure of representational spaces subserving object recognition [44,45,113]. Specifically, RSA estimates the pairwise distance between neural measures associated with processing different stimuli (e.g., images of different objects) and infers, based on such differences, the representational (dis)similarity between the corresponding stimuli. RSA is capable of targeting a broad variety of representations (including conceptual and task-related) in addition to perceptual representations. However, such broadness can also yield underconstrained and ambiguous interpretations of similarity-based results [114]. Relative to RSA, image reconstruction is limited in its immediate scope to visual representations. However, by virtue of such focus, reconstruction can resolve the features underlying the structure of representational spaces by extracting and visualizing the informational content responsible for this structure.

Another class of relevant methods includes noise-based reverse correlation and Bubbles. The former estimates a visual template, aimed at approximating a target representation, from patterns of visual noise (e.g., white noise) added to or supplanting a stimulus image [115,116]. The latter estimates the components of an image (e.g., specific areas or spatial frequency bands) that guide behavioral and/or neural responses [117,118]. These two methods have provided important insights into the visual information that underlies face recognition in both healthy individuals and in those with impaired face processing [119,120]. Similarly to image reconstruction, these methods uncover diagnostic image information underlying visual representations. In particular, noise-based reverse correlation aims to synthesize the appearance of a representation, although, unlike reconstruction, it relies on visual noise as opposed to more complex visual features. Importantly, noise-based reverse correlation, through its use of visual noise, limits the possibility of specific biases associated with reliance on inaccurate or incomplete sets of visual features for reconstruction purposes. However, the substantial number of trials that are necessary to recover the appearance of a single template renders noise-based reverse correlation impractical when applied to a large set of target stimuli and corresponding representations (e.g., dozens or hundreds of individual face representations).

Thus, image reconstruction provides a related but complementary approach relative to current methodological paradigms. Of note, some types of image reconstruction capitalize on the strengths of these different paradigms and, further, they even embed techniques akin to reverse correlation and RSA into image reconstruction [29,30].

attempts require rigorous characterization of individual differences in face representation and their neural correlates [103–105]. We believe that current efforts offer a solid foundation for progress in this regard.

Methodologically, because different reconstruction approaches exhibit different biases, we also note the need to carefully compare their results. The present review takes a step in this direction, although experimental studies that compare different approaches across the same population and stimuli should be able to provide more precise insights. Further, reconstruction would benefit from its comparison with the outcomes of more traditional methods such as noise-based reverse correlation (Box 3).

Although the present review focuses on faces, further evaluation of reconstruction, as applied more generally to natural images and to other categories of interest, will be beneficial. In particular, we note the recent extension of specific face decoding and reconstruction approaches, as reviewed [32], to other types of stimuli such as words [106,107]. Hence, future work will need to consider the promise and the limits of cross-domain reconstruction.

To conclude, image reconstruction provides an informative and comprehensive account of visual face representations that has notable theoretical and practical prospects. This approach can help to answer longstanding questions in the study of face recognition while also guiding similar endeavors in other visual domains.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (to A.N. and A.L.) and by the National Institutes of Health (grant RO1EY027018 to M.B. and D.P.).

References

1. Haxby, J.V. and Gobbini, M.I. (2012) Distributed neural systems for face perception. In *Oxford Handbook of Face Perception* (Calder, A. *et al.*, eds), pp. 93–110, Oxford University Press
2. Freiwald, W. *et al.* (2016) Face processing systems: from neurons to real-world social perception. *Annu. Rev. Neurosci.* 39, 325–346
3. O'Toole, A.J. *et al.* (2018) Face space representations in deep convolutional neural networks. *Trends Cogn. Sci.* 22, 794–809
4. Young, A.W. and Burton, A.M. (2018) Are we face experts? *Trends Cogn. Sci.* 22, 100–110
5. Kriegeskorte, N. *et al.* (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 104, 20600–20605
6. Nestor, A. *et al.* (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9998–10003
7. Anzellotti, S. *et al.* (2013) Decoding representations of face identity that are tolerant to rotation. *Cereb. Cortex* 24, 1988–1995
8. Goesaert, E. and de Beeck, H.P. (2013) Representations of facial identity information in the ventral visual stream investigated with multivoxel pattern analyses. *J. Neurosci.* 33, 8549–8558
9. Natu, V.S. *et al.* (2010) Dissociable neural patterns of facial identity across changes in viewpoint. *J. Cogn. Neurosci.* 22, 1570–1582
10. Vida, M.D. *et al.* (2017) Spatiotemporal dynamics of similarity-based neural representations of facial identity. *Proc. Natl. Acad. Sci. U. S. A.* 114, 388–393
11. Nemrodov, D. *et al.* (2016) The time course of individual face recognition: a pattern analysis of ERP signals. *Neuroimage* 132, 469–476
12. Dobs, K. *et al.* (2019) How face perception unfolds over time. *Nat. Commun.* 10, 1258
13. Ghuman, A.S. *et al.* (2014) Dynamic encoding of face information in the human fusiform gyrus. *Nat. Commun.* 5, 5672
14. Smith, F.W. and Smith, M.L. (2019) Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. *Neuroimage* 195, 261–271
15. Jack, R.E. and Schyns, P.G. (2017) Toward a social psychophysics of face communication. *Annu. Rev. Psychol.* 68, 269–297
16. Schyns, P.G. *et al.* (2020) Revealing the information contents of memory within the stimulus information representation framework. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190705
17. Naselaris, T. *et al.* (2009) Bayesian reconstruction of natural images from human brain activity. *Neuron* 63, 902–915
18. Horikawa, T. and Kamitani, Y. (2017) Generic decoding of seen and imagined objects using hierarchical visual features. *Nat. Commun.* 8, 15037
19. Seeliger, K. *et al.* (2018) Generative adversarial networks for reconstructing natural images from brain activity. *Neuroimage* 181, 775–785
20. Shen, G. *et al.* (2019) Deep image reconstruction from human brain activity. *PLoS Comput. Biol.* 15, e1006633
21. Nishimoto, S. *et al.* (2011) Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* 21, 1641–1646
22. Wen, H. *et al.* (2018) Neural encoding and decoding with deep learning for dynamic natural vision. *Cereb. Cortex* 28, 4136–4160
23. Thirion, B. *et al.* (2006) Inverse retinotopy: inferring the visual content of images from brain activation patterns. *Neuroimage* 33, 1104–1116
24. Miyawaki, Y. *et al.* (2008) Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron* 60, 915–929
25. Schoenmakers, S. *et al.* (2013) Linear reconstruction of perceived images from human brain activity. *Neuroimage* 83, 951–961
26. Du, C. *et al.* (2019) Reconstructing perceived images from human brain activities with bayesian deep multiview learning. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 2310–2323
27. Lescroart, M.D. and Gallant, J.L. (2019) Human scene-selective areas represent 3D configurations of surfaces. *Neuron* 101, 178–192

28. Cowen, A.S. *et al.* (2014) Neural portraits of perception: reconstructing face images from evoked brain activity. *Neuroimage* 94, 12–22
29. Nestor, A. *et al.* (2016) Feature-based face representations and image reconstruction from behavioral and neural data. *Proc. Natl. Acad. Sci. U. S. A.* 113, 416–421
30. Chang, C.H. *et al.* (2017) Memory and perception-based facial image reconstruction. *Sci. Rep.* 7, 6499
31. Zhan, J. *et al.* (2019) Modelling face memory reveals task-generalizable representations. *Nat. Hum. Behav.* 3, 817–826
32. Nemrodov, D. *et al.* (2018) The neural dynamics of facial identity processing: insights from EEG-based pattern analysis and image reconstruction. *eNeuro* 5, e0358-17.2018
33. Nemrodov, D. *et al.* (2019) Multimodal evidence on shape and surface information in individual face processing. *Neuroimage* 184, 813–825
34. Chang, L. and Tsao, D.Y. (2017) The code for facial identity in the primate brain. *Cell* 169, 1013–1028
35. Güçlütürk, Y. *et al.* (2017) Reconstructing perceived faces from brain activations with deep adversarial neural decoding. In *Advances in Neural Information Processing Systems 30* (Guyon, I. *et al*., eds), pp. 4246–4257, Neural Information Processing Systems Foundation
36. VanRullen, R. and Reddy, L. (2019) Reconstructing faces from fMRI patterns using deep generative neural networks. *Commun. Biol.* 2, 193
37. Maurer, D. *et al.* (2002) The many faces of configural processing. *Trends Cogn. Sci.* 6, 255–260
38. Haxby, J.V. *et al.* (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430
39. Valentine, T. *et al.* (2016) Face-space: a unifying concept in face recognition research. *Q. J. Exp. Psychol.* 69, 1996–2019
40. Leopold, D.A. *et al.* (2001) Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat. Neurosci.* 4, 89–94
41. Loffler, G. *et al.* (2005) fMRI evidence for the neural representation of faces. *Nat. Neurosci.* 8, 1386–1390
42. Gao, X. and Wilson, H.R. (2013) The neural representation of face space dimensions. *Neuropsychologia* 51, 1787–1793
43. Nestor, A. *et al.* (2013) Face-space architectures: evidence for the use of independent color-based features. *Psychol. Sci.* 24, 1294–1300
44. Kriegeskorte, N. and Diedrichsen, J. (2019) Peeling the onion of brain representations. *Annu. Rev. Neurosci.* 42, 407–432
45. Kriegeskorte, N. and Douglas, P.K. (2019) Interpreting encoding and decoding models. *Curr. Opin. Neurobiol.* 55, 167–179
46. Livingstone, M. and Hubel, D. (1988) Segregation of form, color, and depth: anatomy, physiology, and perception. *Science* 240, 740–749
47. Bruce, V. and Young, A.W. (2012) *Face Perception*, Psychology Press
48. Russell, R. *et al.* (2007) The utility of surface reflectance for the recognition of upright and inverted faces. *Vis. Res.* 47, 157–165
49. Andrews, T.J. *et al.* (2016) Contributions of feature shapes and surface cues to the recognition and neural representation of facial identity. *Cortex* 83, 280–291
50. Jiang, F. *et al.* (2009) Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience* 163, 1078–1091
51. Cootes, T.F. *et al.* (2001) Active appearance models. *EEE Trans. Pattern Anal. Mach. Intell.* 23, 681–685
52. Quiroga, R.Q. *et al.* (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–1107
53. Rey, H.G. *et al.* (2020) Single neuron coding of identity in the human hippocampal formation. *Curr. Biol.* 30, 1152–1159
54. Rajimehr, R. *et al.* (2009) An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1995–2000
55. Tsao, D.Y. *et al.* (2008) Comparing face patch systems in macaques and humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19513–19518
56. Grill-Spector, K. and Malach, R. (2004) The human visual cortex. *Annu. Rev. Neurosci.* 27, 649–677
57. Guntupalli, J.S. *et al.* (2017) Disentangling the representation of identity from head view along the human face processing pathway. *Cereb. Cortex* 27, 46–53
58. Itier, R.J. and Taylor, M.J. (2002) Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *Neuroimage* 15, 353–372
59. Caharel, S. *et al.* (2009) Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia* 47, 639–643
60. Grossman, S. *et al.* (2019) Convergent evolution of face spaces across human face-selective neuronal groups and deep convolutional networks. *Nat. Commun.* 10, 4934
61. Hill, M.Q. *et al.* (2019) Deep convolutional neural networks in the face of caricature. *Nat. Mach. Intell.* 1, 522–529
62. Grill-Spector, K. *et al.* (2018) The functional neuroanatomy of face perception: from brain measurements to deep neural networks. *Interface Focus* 8, 20180013
63. Yildirim, I. *et al.* (2020) Efficient inverse graphics in biological face processing. *Sci. Adv.* 6, eaax5979
64. Phillips, P.J. *et al.* (2018) Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proc. Natl. Acad. Sci. U. S. A.* 115, 6171–6176
65. Yamins, D.L.K. and DiCarlo, J.J. (2016) Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* 19, 356–365
66. Cichy, R.M. *et al.* (2016) Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* 6, 27755
67. Kietzmann, T.C. *et al.* (2019) Recurrence is required to capture the representational dynamics of the human visual system. *Proc. Natl. Acad. Sci. U. S. A.* 116, 21854–21863
68. Kirby, M. and Sirovich, L. (1990) Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 103–108
69. Dijkstra, N. *et al.* (2019) Shared neural mechanisms of visual perception and imagery. *Trends Cogn. Sci.* 23, 423–434
70. Breedlove, J.L. *et al.* (2020) Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol.* Published online April 30, 2020, https://doi.org/10.1016/j.cub.2020.04.014
71. Linde-Domingo, J. *et al.* (2019) Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat. Commun.* 10, 179
72. Lee, H. and Kuhl, B.A. (2016) Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *J. Neurosci.* 36, 6069–6082
73. Cabeza, R. *et al.* (2008) The parietal cortex and episodic memory: an attentional account. *Nat. Rev. Neurosci.* 9, 613–625
74. Graham, K.S. *et al.* (2010) Going beyond LTM in the MTL: a synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia* 48, 831–853
75. Blauch, N.M. *et al.* (2019) Computational insights into human expertise for familiar and unfamiliar face recognition. *PsyArXiv* Published online October 24, 2019. http://dx.doi.org/10.31234/osf.io/bv5mp
76. Geskin, J. and Behrmann, M. (2018) Congenital prosopagnosia without object agnosia? A literature review. *Cogn. Neuropsychol.* 35, 4–54
77. Barton, J. *et al.* (2016) Prosopagnosia: current perspectives. *Eye Brain* 8, 165–175
78. Richoz, A.R. *et al.* (2015) Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex* 65, 50–64
79. Rossion, B. *et al.* (2003) A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126, 2381–2395
80. Ekman, P. and Friesen, W.V. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press
81. Benitez-Quiroz, C.F. *et al.* (2018) Facial color is an efficient mechanism to visually transmit emotion. *Proc. Natl. Acad. Sci. U. S. A.* 115, 3581–3586

82. Du, S. *et al.* (2014) Compound facial expressions of emotion. *Proc. Natl. Acad. Sci. U. S. A.* 111, E1454–E1462
83. Chang, C. *et al.* (2019) Image reconstruction reveals the impact of aging on face perception. *bioRxiv* Published online December 20, 2019, https://doi.org/10.1101/2019.12.20.883215
84. Boutet, I. *et al.* (2015) On the particular vulnerability of face recognition to aging: a review of three hypotheses. *Front. Psychol.* 6, 1139
85. Lee, Y. *et al.* (2014) Broadly tuned face representation in older adults assessed by categorical perception. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1060–1071
86. Konar, Y. *et al.* (2013) Effects of aging on face identification and holistic face processing. *Vis. Res.* 88, 38–46
87. Lee, Y. *et al.* (2011) Face processing changes in normal aging revealed by fMRI adaptation. *J. Cogn. Neurosci.* 23, 3433–3447
88. Gilad, S. *et al.* (2009) Role of ordinal contrast relationships in face encoding. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5353–5358
89. Creighton, S.E. *et al.* (2019) Classification images characterize age-related deficits in face discrimination. *Vis. Res.* 157, 97–104
90. Whitney, D. and Yamanashi Leib, A. (2018) Ensemble perception. *Annu. Rev. Psychol.* 69, 105–129
91. Haberman, J. and Whitney, D. (2007) Rapid extraction of mean emotion and gender from sets of faces. *Curr. Biol.* 17, 751–753
92. Neumann, M.F. *et al.* (2013) Viewers extract mean and individual identity from sets of famous faces. *Cognition* 128, 56–63
93. De Fockert, J. and Wolfenstein, C. (2009) Rapid extraction of mean identity from sets of faces. *Q. J. Exp. Psychol.* 62, 1716–1722
94. Roberts, T. *et al.* (2019) Elucidating the neural representation and the processing dynamics of face ensembles. *J. Neurosci.* 39, 7737–7747
95. Krakauer, J.W. *et al.* (2017) Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93, 480–490
96. Dalrymple, K.A. *et al.* (2014) Spontaneous perceptual facial distortions correlate with ventral occipitotemporal activity. *Neuropsychologia* 59, 179–191
97. Daros, A.R. *et al.* (2013) Facial emotion recognition in borderline personality disorder. *Psychol. Med.* 43, 1953–1963
98. Tsoi, D.T. *et al.* (2008) Is facial emotion recognition impairment in schizophrenia identical for different emotions? A signal detection analysis. *Schizophr. Res.* 99, 263–269
99. Uchiyama, M. *et al.* (2012) Pareidolias: complex visual illusions in dementia with Lewy bodies. *Brain* 135, 2458–2469
100. Lee, A.C.H. *et al.* (2006) Differentiating the roles of the hippocampus and perirhinal cortex in processes beyond long-term declarative memory: a double dissociation in dementia. *J. Neurosci.* 26, 5198–5203
101. Lavallée, M.M. *et al.* (2016) A qualitative impairment in face perception in Alzheimer's disease: evidence from a reduced face inversion effect. *J. Alzheimers Dis.* 51, 1225–1236
102. Jayaram, V. *et al.* (2016) Transfer learning in brain–computer interfaces. *IEEE Comput. Intell. Mag.* 11, 20–31
103. McGugin, R.W. and Gauthier, I. (2016) The reliability of individual differences in face-selective responses in the fusiform gyrus and their relation to face recognition ability. *Brain Imaging Behav.* 10, 707–718
104. Wilmer, J.B. (2017) Individual differences in face recognition: a decade of discovery. *Curr. Dir. Psychol. Sci.* 26, 225–230
105. Peterson, M.F. *et al.* (2016) Individual differences in face-looking behavior generalize from the lab to the world. *J. Vis.* 16, 1–18
106. Ling, S. *et al.* (2019) How are visual words represented? Insights from EEG-based visual word decoding, feature derivation and image reconstruction. *Hum. Brain Mapp.* 40, 5056–5068
107. Nemrodov, D. *et al.* (2020) A multivariate investigation of visual word, face, and ensemble processing: perspectives from EEG-based decoding and feature selection. *Psychophysiology* 57, e13511
108. Ramírez, F.M. (2018) Orientation encoding and viewpoint invariance in face recognition: inferring neural properties from large-scale signals. *Neuroscientist* 24, 582–608
109. Stanley, G.B. *et al.* (1999) Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci.* 19, 8036–8042
110. deCharms, R.C. and Zador, A. (2000) Neural representation and the neural code. *Annu. Rev. Neurosci.* 23, 613–647
111. Valentine, T. *et al.* (2016) Face-space: a unifying concept in face recognition research. *Q. J. Exp. Psychol.* 69, 1996–2019
112. Ross, D.A. *et al.* (2014) Not just the norm: exemplar-based models also predict face aftereffects. *Psychon. Bull. Rev.* 21, 47–70
113. Cichy, R.M. *et al.* (2014) Resolving human object recognition in space and time. *Nat. Neurosci.* 17, 455–462
114. Medin, D.L. *et al.* (1993) Respects for similarity. *Psychol. Rev.* 100, 254–278
115. Gosselin, F. and Schyns, P.G. (2003) Superstitious perceptions reveal properties of internal representations. *Psychol. Sci.* 14, 505–509
116. Nestor, A. *et al.* (2013) Internal representations for face detection: an application of noise-based image classification to BOLD responses. *Hum. Brain Mapp.* 34, 3101–3115
117. Schyns, P.G. *et al.* (2007) Dynamics of visual information integration in the brain for categorizing facial expressions. *Curr. Biol.* 17, 1580–1585
118. Zhan, J. *et al.* (2019) Dynamic construction of reduced representations in the brain for perceptual decision behavior. *Curr. Biol.* 29, 319–326
119. Adolphs, R. *et al.* (2005) A mechanism for impaired fear recognition after amygdala damage. *Nature* 433, 68–72
120. Tardif, J. *et al.* (2019) Use of face information varies systematically from developmental prosopagnosics to super-recognizers. *Psychol. Sci.* 30, 300–308