



# FaReT: A free and open-source toolkit of three-dimensional models and software to study face perception

Jason Hays<sup>1</sup> · Claudia Wong<sup>1</sup> · Fabian A. Soto<sup>1</sup>

Published online: 9 June 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

A problem in the study of face perception is that results can be confounded by poor stimulus control. Ideally, experiments should precisely manipulate facial features under study and tightly control irrelevant features. Software for 3D face modeling provides such control, but there is a lack of free and open source alternatives specifically created for face perception research. Here, we provide such tools by expanding the open-source software MakeHuman. We present a database of 27 identity models and six expression pose models (sadness, anger, happiness, disgust, fear, and surprise), together with software to manipulate the models in ways that are common in the face perception literature, allowing researchers to: (1) create a sequence of renders from interpolations between two or more 3D models (differing in identity, expression, and/or pose), resulting in a “morphing” sequence; (2) create renders by extrapolation in a direction of face space, obtaining 3D “anti-faces” and caricatures; (3) obtain videos of dynamic faces from rendered images; (4) obtain average face models; (5) standardize a set of models so that they differ only in selected facial shape features, and (6) communicate with experiment software (e.g., PsychoPy) to render faces dynamically online. These tools vastly improve both the speed at which face stimuli can be produced and the level of control that researchers have over face stimuli. We validate the face model database and software tools through a small study on human perceptual judgments of stimuli produced with the toolkit.

**Keywords** Face database · Face morphing · Face identity · Face expression · Computer-generated faces · Dynamic faces

Psychophysics is the quantitative study of the relation between physical stimuli and perception (Kingdom and Prins, 2016). As such, its success is determined by our ability to precisely manipulate and measure physical stimuli. This is a relatively straightforward task when the focus is on low-level stimulus features, such as brightness, color, contrast, or spatial frequency information, among others. However, the task becomes more complex when one is interested in the stimulus features that affect our perception of complex objects and scenes. We sorely lack methods to produce naturalistic stimuli in a controlled manner, or to measure their relevant properties.

In vision science, faces are one class of complex objects for which we have made important advances. An advantage of faces over other complex objects is that images of different

faces can be aligned in their main features (e.g., nose, eyes, mouth). This has enabled the adoption of morphing techniques as a way to manipulate complex face dimensions in a relatively controlled manner (e.g., Steyvers, 1999), the application of traditional psychophysical techniques to estimate templates used to classify faces, such as reverse correlation or bubbles (e.g., Gosselin and Schyns, 2001; Macke and Wichmann, 2010; Mangini and Biederman, 2004; Schyns et al., 2002; Soto, 2019), and the application of multivariate statistics to study the distribution of naturally varying faces, through principal component analysis (PCA) and related techniques (Calder et al., 2001; O’Toole et al., 1993; Turk and Pentland, 1991). For the past two decades, research on face perception has been strongly supported by the availability of such techniques and a vast number of face photograph databases (e.g., Dailey et al., 2001; Ebner et al., 2010; Lucey et al., 2010; Lundqvist et al., 1998; Ma et al., 2015; Strohminger et al., 2016).

However, naturalistic face databases and two-dimensional image manipulation techniques are sometimes insufficient to obtain the tight stimulus control required to answer certain

---

✉ Fabian A. Soto  
fasoto@fiu.edu

<sup>1</sup> Department of Psychology, Florida International University, Miami, FL, USA

questions in face perception research. For example, imagine an experiment aimed at determining whether people perceive males as angrier than females (e.g., Aguado et al., 2009; Bayet et al., 2015; Becker et al., 2007). An experimenter obtains photographs of male and female actors showing angry expressions, and it is observed that people are faster at recognizing anger in males than in females. Unfortunately, such results could have nothing to do with face perception, and everything to do with the production of expressions by the actors. If males simply produce stronger expressions of anger, then the observed results are not due to perception, but to the specific stimuli used. The exact poses by actors and the strength of their expression are difficult to control using naturalistic photographs and morphing techniques. It would be desirable to have a way to obtain the exact same expression pose from faces with different identities and sexes. It would also be desirable to easily standardize a set of faces with regard to features that are not of interest in a particular study. Examples include texture and coloration of the face, shape of external features like ears and neck, eye color, and facial hair. Other studies would want to precisely manipulate such features rather than simply control them. For example, one might want to answer the question of whether neck width, face color, or facial hair, all of which differ between sexes, is what facilitates perception of anger in male faces. Such control can possibly be achieved using two-dimensional image manipulation, but the process is painstaking and has traditionally involved some degree of manual image manipulation.

Morphing sequences are also used as a way to obtain dynamic stimuli. For example, some studies on face expression have presented morphing sequences going from neutral to some target expression, by concatenating them in a video representing dynamic emotional expressions. However, this approach has been criticized on the basis that such morphing sequences might not reflect the changes in face shape that would occur naturally during dynamic expression (Bernstein and Yovel, 2015; Roesch et al., 2011).

An additional limitation imposed by techniques that manipulate two-dimensional images is that they tend to be obscure regarding the specific properties of the physical object that are being modified and the extent of such changes. Take the example of face morphing. There are several different algorithms that can be used to produce morphed sequences. They all first require templates to be manually created for each two-dimensional image that will be included in the morph, which characterize the position of landmark features of the depicted objects. Such templates are used to apply a complex nonlinear transformation from one image to another, which is summarized by a single number representing “percentage of change”. Both the morphing transformation space and the changes produced within it are hard to interpret. Ideally, one would want to have the ability to more precisely manipulate specific shape features of the relevant object, and to describe

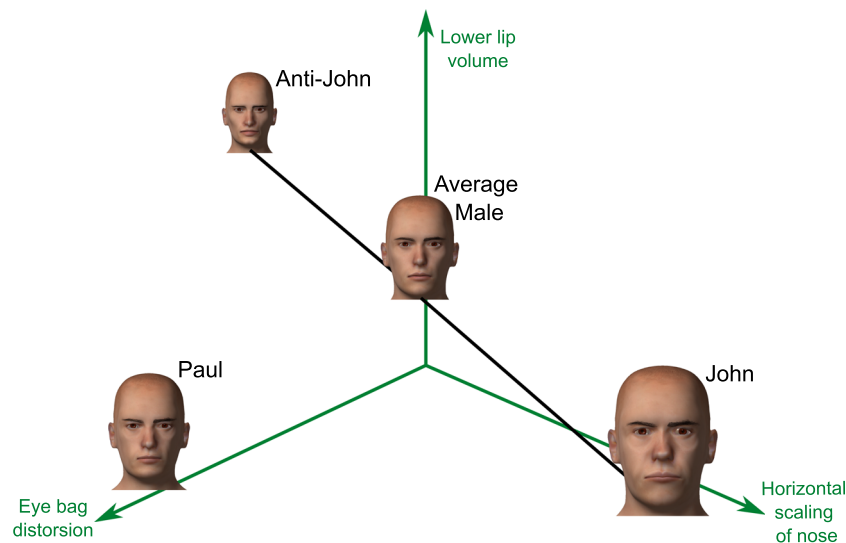
transformations in terms of such interpretable shape features. Similarly, results from PCA and related multivariate techniques are often difficult to interpret, as the analysis is performed in the space of image pixels rather than in a space composed of higher-level face features. In addition, the obtained statistical description applies to face images obtained from a specific viewpoint (usually frontal).

The use of three-dimensional computer graphics face models provides researchers with more precise control over face stimuli. Indeed, some researchers in the visual neuroscience community conclude that the level of control necessary to answer some important standing questions can only be achieved through the use of computer-generated face stimuli (Anzellotti and Caramazza, 2014). However, such models are usually described by thousands of parameters that are not more interpretable than image pixels. For this reason, it would be desirable to have software able to create and manipulate realistic three-dimensional face models that are described through a smaller number of parameters that are easily interpretable by researchers.

## Face research toolkit (FaReT)

The use of three-dimensional models to generate stimuli for face research has the advantage of providing strong stimulus control without losing a naturalistic appearance. However, such models are usually described as meshes with thousands of parameters, making them just as difficult to interpret as two-dimensional images described in terms of pixel luminance. In addition, such complexity is usually encapsulated by intentional obscurity: the available face databases and relevant morphing software are often proprietary and closed-source, limiting the community’s ability to access, expand, and comprehend the resources. Thankfully, MakeHuman (<http://www.makehumancommunity.org/>; Bastioni et al., 2008; Ceipidor et al., 2008) is a free and open-source computer graphics software that allows users to create three-dimensional face models via description through a relatively small number of easily interpretable and sensibly named parameters (nostrils angle, cheeks outer volume, etc.). In MakeHuman, any face model can be described as a point in a multidimensional shape space, with each shape parameter representing one dimension, as depicted in Fig. 1. The software uses smart algorithms to convert changes in this shape space into changes in the higher-dimensional space of a face mesh.

Our goal is to enhance MakeHuman’s capabilities in order to facilitate its use by face perception researchers. With this goal in mind, we have developed the Face Research Toolkit, or FaReT, which is composed of several Python plugins that add utilities to MakeHuman. These plugins were developed to provide face researchers transitioning from image morphs/



**Fig. 1** Schematic representation of how face identity models are generated in MakeHuman. Each model is represented as a point in a multidimensional shape space, where a dimension represents a nameable shape feature (e.g., eye bag distortion, lower lip volume,

horizontal scaling of nose). Different models differ in one or more shape dimensions. If a model (e.g., John) is connected to the average model through a line, models on the other side of that line can be considered anti-identity models (e.g., anti-John)

manipulations with functionality that they would appreciate, being analogous to manipulations that they already use but providing much more control and interpretability. In addition, FaReT includes a database of 27 identity models and 6 expression pose models (sadness, anger, happiness, disgust, fear, and surprise), so researchers can easily get started generating face stimuli using our plugins.

## MakeHuman

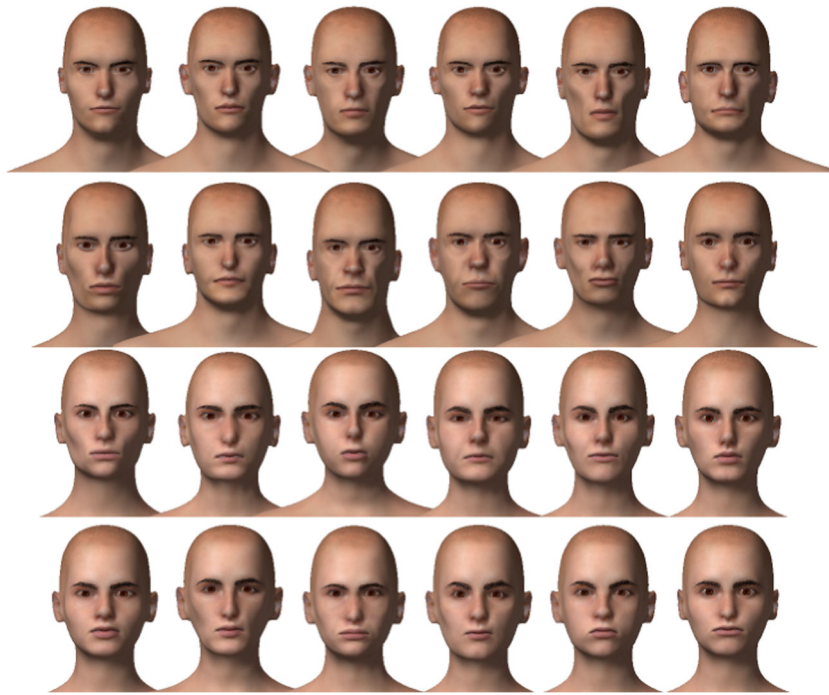
Originally, MakeHuman was intended to help game designers and animators easily create and rig human models, even going so far as to supply a base set of skin textures, clothing assets, and several skeletons with varying complexity. The MakeHuman developers managed to reduce thousands of parameters (the positions of every vertex of a human body mesh, which consists of exclusively quadrilateral polygons) to dozens that are identified through sensible, easily interpreted names (nostrils angle, cheek outer volume, etc.), while also adding safeguards to help keep the mesh's vertices smooth (Bastioni et al., 2008).

MakeHuman allows one to easily change common face parameters like sex, age, ethnicity, and level of face fat of the face models, by manipulating a single slider. In theory, any identity can be reproduced through the software, and it is also possible to create completely new realistic faces. Interaction with the software is through an intuitive graphical user interface that does not require prior knowledge or training. Skin texture models can be obtained from MakeHuman or developed independently, and they can be placed in any face model.

Similarly, an expression pose model can also be developed independently and applied to any face model in MakeHuman. The software itself comes with several pose models, and others can be obtained from the active online community. In addition, we have developed our own set of standardized basic emotional expressions (see below), based on pictures from one actor in the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist et al., 1998).

Just as with identity, MakeHuman allows one to easily develop a pose model for any facial expression through its *expression mixer* plugin. The plugin involves a number of expression parameters that have easily interpretable names, such as “left eye down”, “right inner brow up”, and so on. Each parameter has its own slider in MakeHuman’s graphical interface, which facilitates the creation of new pose models without any prior training. The MPEG-4 standard (Pandzic and Forchheimer, 2002) influenced the development of MakeHuman’s expression mixer in the early years (Ceipidor et al., 2008), and thus many of MakeHuman’s expression parameters are equivalent to facial animation parameters (FAPs) from that widely used standard. In the behavioral sciences, a more common standard used to generate expressions for face research is the facial action coding system (FACS; Ekman and Friesen, 1975), which is related to the MPEG-4 system (Pandzic and Forchheimer, 2002). Researchers who are interested in using the traditional FACS system to generate expression models can do so in MakeHuman through the FACSHuman plugin (Gilbert et al., 2018; <https://github.com/montybot/FACSHuman>).

Models created with MakeHuman can be exported in formats that are easy to open and use in other 3D modeling and game development software, such as Blender ([Springer](https://www.</a></p>
</div>
<div data-bbox=)

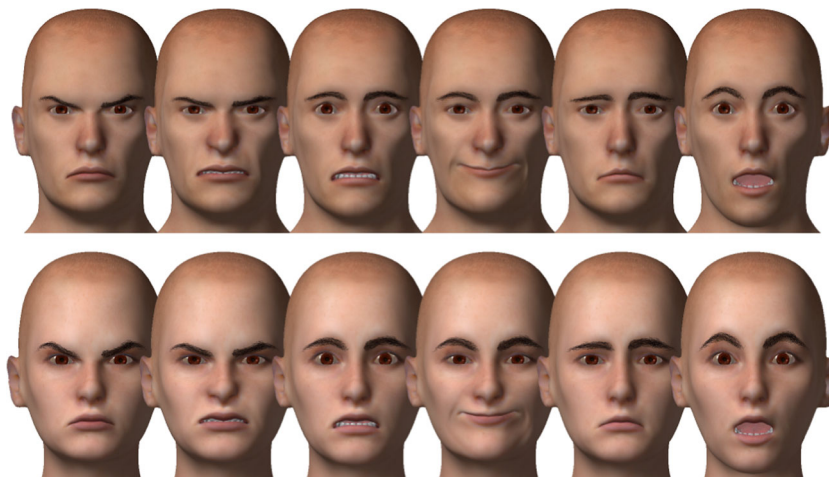


**Fig. 2** Renders of the 24 identity models included with FaReT. The top two rows are the male identities and the bottom two rows are the female identities

[blender.org/](https://blender.org/)). This greatly expands the types of research that can be performed using the 3D models created with MakeHuman. Interactive games, virtual reality (VR) videos, and motion capture are all possibilities that are relatively accessible to an interested researcher with enough technical knowledge, facilitated by the fast and easy creation of 3D face models in MakeHuman.

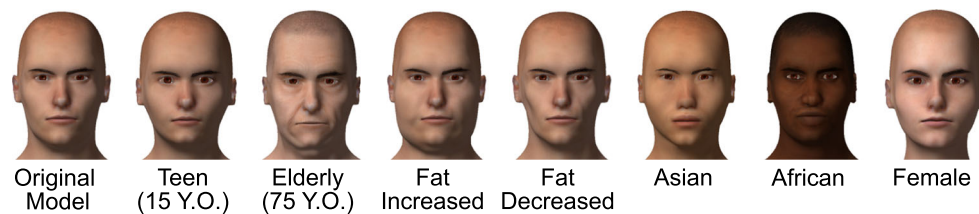
The functionality provided by MakeHuman can be obtained from other software, such as FaceGen Modeller expanded with the FACSGen plugin (Roesch et al., 2011). Stimuli

created using FaceGen have been relatively popular among face researchers in recent years (e.g., Ho et al., 2018; Lamer et al., 2017; Oosterhof and Todorov, 2008; Soto, 2019; Soto and Ashby, 2015, 2019; Thorstenson et al., 2019; Uddenberg and Scholl, 2018). However, an important advantage of MakeHuman is that it is open-source software. This means that it can be downloaded and used absolutely free of charge by the scientific community. In addition, the base code for MakeHuman is available for examination and expansion by interested researchers, so any features that the software lacks



**Fig. 3** Renders of the six expression pose models originally included with FaReT, applied to two different identity models. Each row contains one identity; the top row is a male identity and the bottom is a

female identity. Each column contains a separate expression: anger, disgust, fear, happiness, sadness, and surprise



**Fig. 4** Example of how one of the models in our database (CM01: male, 25 years old, Caucasian, average face fat) can easily be used to generate novel models with a different age, level of face fat, race, or sex. Each one

can be added by the scientific community. Plugins and other add-ons like those provided by FaReT are completely available for anyone not only to use, but also to expand and maintain, even if the original developers discontinue maintenance. Such plugins are written using the Python programming language, which has in recent years become the standard in several areas of psychology and neuroscience.

In contrast, the use of commercial software like FaceGen Modeller can stifle scientific development. For example, there was a period of several years when the incredibly useful FACSGen was no longer being actively developed. If FACSGen had been developed on top of open-source software like MakeHuman, interested researchers could have taken the project into their own hands for maintenance and updating. Although FACSGen is now being developed and maintained again, a license must be purchased both for the plugin and for the base FACEGen software. Thus, only researchers with financial resources can use this setup.

### FaReT's database of three-dimensional face models

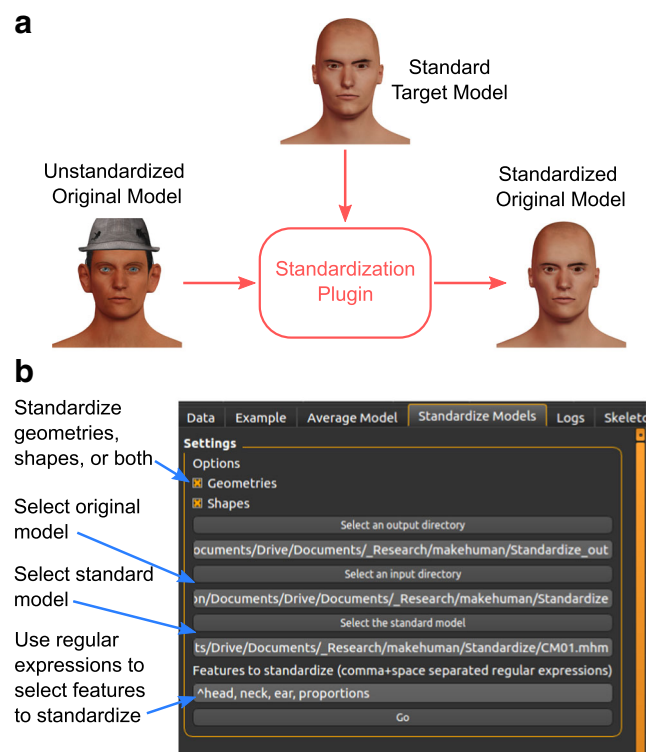
To help us establish the usefulness of the toolbox, we created a database of 24 identities. The identity database consists of 12 male and 12 female models. All 24 models are inspired by photographs of real people's faces; however, the models went through several iterations of modifications to avoid matching the actual identities, while also ensuring that the faces were realistic. The final versions were judged by all three authors as realistic. See Fig. 2 for a frontal view of the final models' faces.

To add further utility to the database, we also included standardized facial expressions. Using MakeHuman's expression mixer, we established six poses for expressions that convey six basic emotions (Ekman, 1999): happy, sad, anger, fear, surprise, and disgust. With the exception of the happy expression, we based our expressions on a single actor in the KDEF database (Lundqvist et al., 1998). The original expression of happy was judged (by the authors) to be somewhat off-putting when applied to some faces, so we manually adjusted it to make it more generally applicable. By generating expression pose files, we can systematically apply the same expression to any identity we created in MakeHuman. See Fig. 3 for a

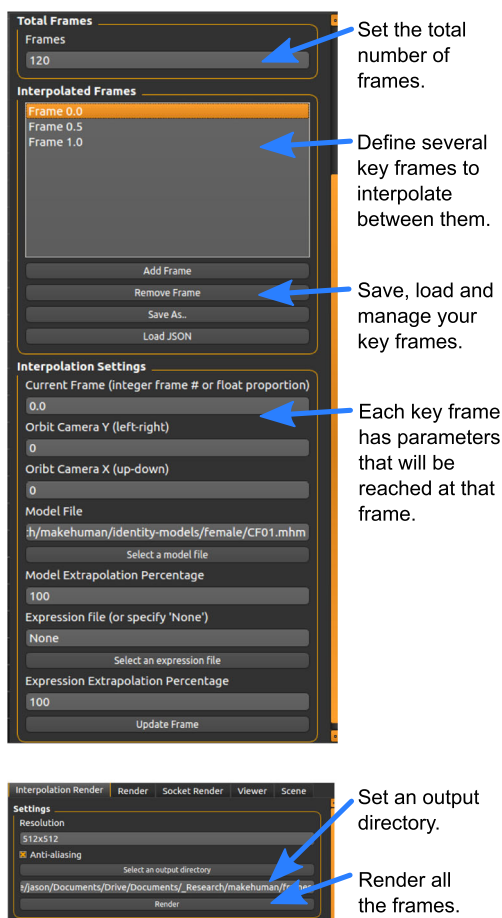
of these models took only seconds to generate in MakeHuman by manipulating key sliders in the GUI (named Age, Head fat, African, Asian, Caucasian, and Gender)

frontal view of two identities with the same expression poses applied. This database was the basis for testing the tools we added to MakeHuman through its Python plugin system.

An advantage of this small database of three-dimensional models is that it can be easily expanded to a much larger database using MakeHuman, including models that are completely different from the originals in some important face category or dimension. Figure 4 shows an example. Here, we have taken a single identity model from our database and used it to easily create seven novel variations. The original model is a 25-year-old Caucasian male with average face fat. By manipulating a single slider in MakeHuman, and in some cases



**Fig. 5** Standardization plugin. Panel (a) shows a schematic representation of the functionality of the standardization plugin. An original unstandardized model is provided together with a target standard model. The plugin outputs a standardized original model. Panel (b) shows the inputs and options included in the plugin's graphical user interface

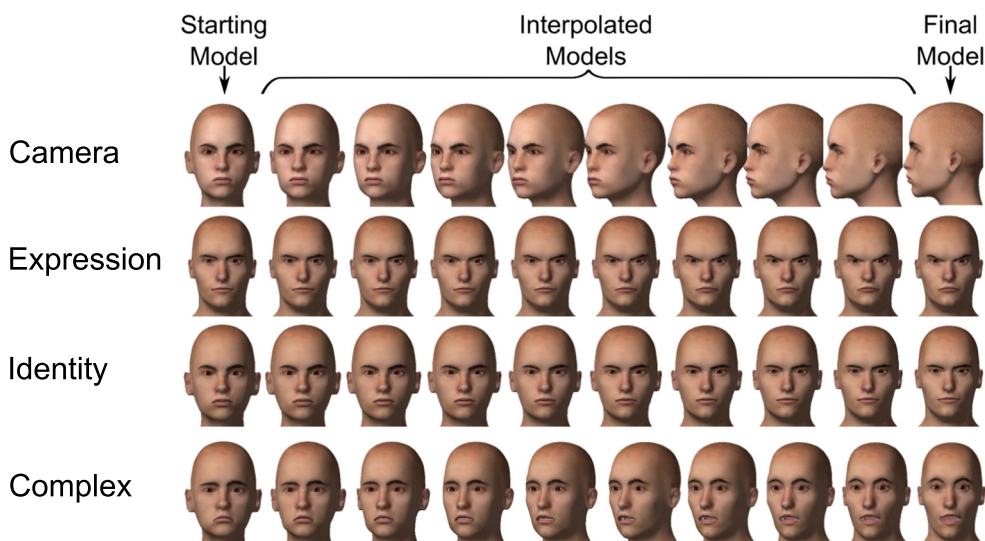


**Fig. 6** Inputs and options included in the graphical user interface of the interpolation render plugin

selecting a different skin model (MakeHuman includes standard skin models for different combinations of age, sex, and race), it is possible to create the following versions of the model: teenager, elderly, fat, skinny, Asian, African, and female. Each one of these models took only seconds to generate, which means that in practice, at least 192 models can be generated from our database, tailored to the goals of each research project.

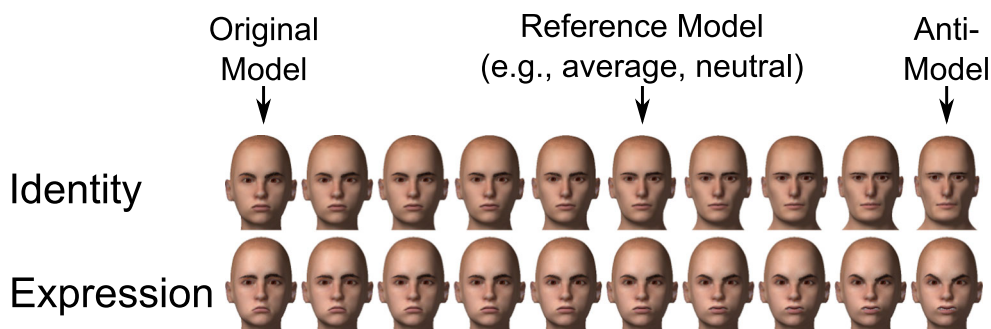
### FaReT plugins

**Averaging and standardization of three-dimensional face models** Taking advantage of MakeHuman's ability to accept Python plugins that extend its usefulness, we designed several such plugins to assist researchers who are interested in fully controlling their stimuli during face perception experiments. Conveniently, MakeHuman creates human-shaped models given a set of simple, named, numerical parameters. Because researchers often need to average faces (i.e., for studies on norm-based face encoding; reviewed by Rhodes 2017; Webster and MacLeod 2011), we created a plugin to help users average the identities generated in MakeHuman by simply obtaining the numerical average of each parameter. The end-user can specify a set of many identity models within a directory and its sub-directory to produce an average identity. In addition to the mean, the plugin also collects information about each parameter's variability, which allows new faces to be generated roughly within the space of the models being averaged.



**Fig. 7** Examples of linear model interpolation using FaReT. Each row shows a different interpolating morph. Each row shows interpolation of: the camera position, expression, identity, each of them simultaneously. The Camera morph is rotating the camera by 90 degrees. The Expression morph is interpolating a male identity between a neutral expression and an angry expression. The Identity morph is interpolating from a female

identity to a male identity: note that because skin tone settings contain Boolean graphical shader settings, the skin tone is not interpolated in the current version of FaReT. The Complex morph is changing a female identity expressing sadness to a male identity expressing surprise while rotating the camera to and from 30 degrees



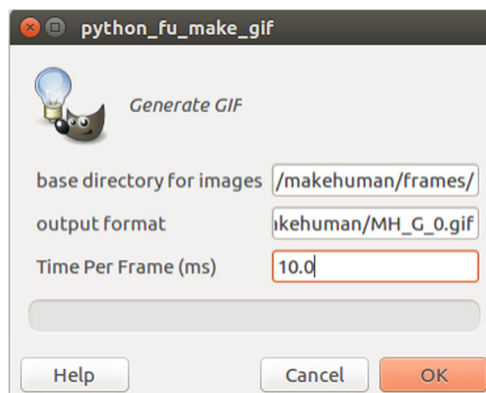
**Fig. 8** Examples of linear extrapolation using FaReT. Depending on the two models chosen, the plugin makes possible morphing to anti-identities and anti-expressions. Creation of caricatures is also available (not shown)

Researchers might be interested in creating face stimuli that differ in a number of chosen features but have standard values in all other attributes. This allows one to make conclusions about the effect of the selected features on perceptual processing, while easily controlling for the possible confounding effect of the nuisance features. To facilitate this process, we developed a standardization plugin. Figure 5a shows a schematic representation of how the standardization process works. The plugin requires as input an original unstandardized model together with a target standard model, and provides as output the original model standardized in selected features. In MakeHuman, there are two types of features: shapes and geometries. The shape parameters only affect the mesh/wireframe of the human model. Geometries are like accessories for the wireframe: skin, teeth, eyebrows, hair, etc. (anything added onto the wireframe). Therefore, we developed a plugin that can standardize one or both types of features across the identity models. When standardizing the shape features, any of the named parameters used to produce the body's wireframe can be standardized (i.e., age, sex, ovalness, etc.). Users can specify comma-space-separated parameter names through Perl-style regular expressions to select the names; i.e., “head” can select anything that contains the ordered letters “head” (forehead-bulge, head-size, etc.), but “^head” can only select parameter names that begin with “head” (i.e., head-size, but not forehead-bulge). As an example of formatting, the default shape standardization setting is “^head, neck, ear, proportions”.

The simpler of the two categories of features is geometries. The geometries of every identity in the specified directory will all be set to match the geometries of the specified standard model, which is selected at the start of the standardization process: meaning that the teeth, eyes, skin tone, etc., will all be set to match the specified model. This is important to consider when using separate skin models for men and women (differences in skin tone and contrast are important cues for sex classification; see Russell 2003, 2009), and simply means that stimuli that require distinct geometries (including skin texture/color models) should be standardized separately.

With our standardization plugin, it is possible to standardize shapes, geometries, or both at the same time. The standardized models are outputted to a separate directory, which the user specifies, to avoid overriding the originals. Figure 5b describes the options offered through the graphical user interface of the standardization plugin.

**“Morphing” in three-dimensional shape space** Morphing is an incredibly important tool in face perception research, but the process to create morphed faces is slow and painstaking, and the resulting sequence is difficult to interpret in terms of meaningful shape parameters. Morphing pixels within images is a nonlinear operation that requires the creation of templates through the manual placement of several nodes to outline each important feature (mouth, eyes, nose, head outline, etc.) for each view, for each expression, and for each face (whose placement is also each subject to human error). Obtaining high-quality morphs usually requires the manual placement of 150+ nodes in each face image (e.g., using Psychomorph or its web-based version, WebMorph). Such templates are view-specific, in that if they are created for a given face view (typically front view), then morphing the same face from a different viewpoint will require the creation of a new, independent template.



**Fig. 9** Graphical user interface of the GIMP plugin used to create high-quality GIF animations from MakeHuman renders

**Table 1** Emotion ratings provided by participants in Study 1

	N	Fear		Sadness		Anger		Happiness		Disgust		Surprise	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Fear	75	<u>5.441</u>	1.299	3.689	1.519	2.183	1.244	1.469	0.660	3.763	1.632	4.847	1.518
Sadness	75	3.779	1.632	<u>5.683</u>	1.308	2.001	1.107	1.375	0.647	2.394	1.409	1.753	1.088
Anger	73	2.025	1.304	2.238	1.372	<u>6.229</u>	0.956	1.258	0.488	3.846	1.808	1.656	0.928
Happiness	76	1.641	0.939	1.691	1.031	1.567	0.792	<u>4.854</u>	1.250	1.723	1.093	1.852	1.172
Disgust	75	2.406	1.557	2.411	1.471	<b>6.178</b>	1.015	1.288	0.660	<u>4.472</u>	1.538	1.768	1.125
Surprise	75	4.417	1.603	2.308	1.350	1.931	1.101	2.028	1.140	2.705	1.557	<u>6.404</u>	0.915

Each row represents a different expression model shown to the participants. Columns represent different expression ratings provided by the participants for the displayed face stimuli. For each identity that was morphed to a specific emotional expression, the participants provided ratings for six emotions on a scale from 1 to 7. For each expression model, the rating corresponding to the intended expression is underlined, and the highest-rated expression is in bold. Only the disgust stimulus was not rated most highly in its corresponding emotion, but confusion between anger and disgust is a common finding in the literature using naturalistic stimuli (see Discussion)

Through MakeHuman, however, we have full control of the models' parameters and can use the line connecting two identity models in shape space (see black line in Fig. 1) to interpolate between them (e.g., any point in the line segment going from John to the average model in Fig. 1) or extrapolate along the line beyond each of them (e.g., the line segment extrapolating toward anti-John in Fig. 1). This reduces the workload required to create image morphs while also increasing the accuracy of the morphs by removing the process of placing nodes altogether. Because MakeHuman models are three-dimensional, any linear interpolation from one face identity or expression to another can be rendered from any viewpoint and in any illumination condition. Finally, because the interpolation is performed in the multidimensional shape space of MakeHuman parameters, the transformation is

interpretable in terms of relative changes in such intuitively named parameters.

Linear interpolation and extrapolation of models is easily achieved in FaReT through the interpolation render plugin. Figure 6 describes the inputs and options included in the graphical user interface of this plugin. The output of the plugin is a sequence of rendered models that can be used to create animated face stimuli, as explained later.

**Linear interpolation in 3D shape space** The interpolation plugin achieves something similar to “morphing” in MakeHuman, by creating a sequence of renders from linear interpolation between two or more 3D models. In addition to the face identity parameters, it is possible to interpolate between expression pose models, orbital camera positions, or any simultaneous combination of those (see Fig. 7).

There are several advantages of this procedure over traditional image morphing. First, the transformation is interpretable in the space of shape parameters. Second, the procedure allows precise control of what parameters are changed, so that “partial morphs” in which only some shape features are changed can be easily achieved. Third, the template creation process can be bypassed while maintaining complete control over the specific, well-defined parameters being changed as well as the number of frames to render. Fourth, as indicated above, interpolation can be performed using not only shape parameters, but also simultaneously expression pose and camera position (see bottom row of Fig. 7).

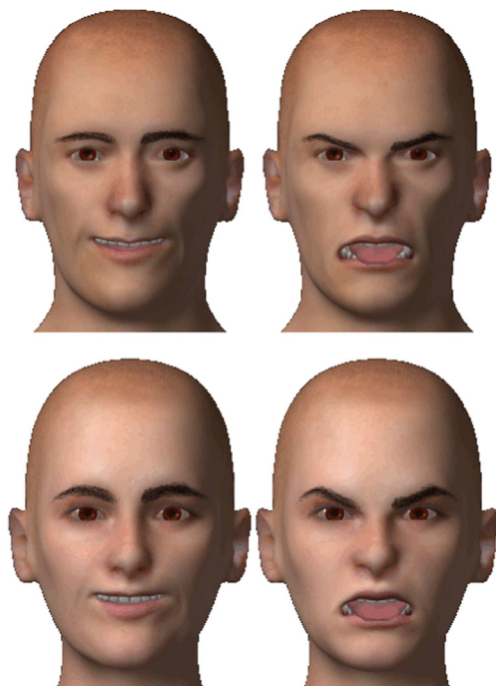
**Linear extrapolation in 3D shape space** During interpolation, the shape parameters of any two identity models, or the pose parameters of any two expressions, were used to create a directional vector from one point in parameter space to the other. The plugin also includes the ability to extrapolate along that vector, beyond the original two models. Extrapolation can be

**Table 2** Unusualness ratings provided by participants in both studies

Unusualness	$\bar{x}$	SD
<i>Study 1</i>		
Fear	4.23	0.71
Sadness	2.94	0.75
Anger	4.43	0.71
Happiness	4.58	0.76
Disgust	4.63	0.57
Surprise	3.90	0.70
<i>Study 2</i>		
Disgust (Tongue Out)	3.89	0.29
Happiness (Teeth)	2.79	0.48

Each row represents a different expression model shown to the participants. Unusualness was rated on a Likert scale from 1 to 7, with higher values representing higher unusualness





**Fig. 10** Renders of the two new expression pose models created for Study 2, applied to two different identity models. Each row contains one identity; the top row is a male identity and the bottom is a female identity. Each column contains a separate expression: open-mouth happiness and tongue-out disgust

applied using any two models, but if a directional vector is defined from an average face to a target identity, then extrapolation can be used to create caricatured identities and expressions by applying the vector even after reaching the target identity’s parameters. Furthermore, by reversing the direction of the vector, extrapolation can be used to create anti-identities and anti-expressions (see Fig. 1). Both caricatures (e.g., Byatt and Rhodes, 1998; Lee et al., 2000) and anti-faces (e.g., Burton et al., 2015; Cook et al., 2011; Leopold et al., 2001; Rhodes and Jeffery, 2006; Skinner and Benton, 2010) have proven to be important stimulus types for examining face encoding in the human visual system. See Fig. 8 for linear extrapolation examples.

**Obtaining videos of dynamic faces** Face perception researchers have recently become increasingly interested in using dynamic face stimuli in their research, particularly in the study of emotional expression (for reviews, see Bernstein and Yovel, 2015; Duchaine and Yovel, 2015; Lander and Butcher, 2015). A common practice is to create sequences of morphed images from a neutral expression to some target expression, and concatenate such images into an animated video. A similar process is possible using FaReT. Our plugins automate the linear interpolation/extrapolation process and output a sequence of images that can be used to create videos or animated GIFs (see Fig. 7), which can be presented as experimental stimuli. Because of the high-quality GIFs that the GNU Image Manipulation Program (GIMP; <https://www.gimp.org/>) exports (by surpassing the standard 256 color limit for GIFs), we have also created a python plugin for GIMP 2.8 to help in transforming folders of MakeHuman renders to GIFs. Figure 9 shows the graphical user interface developed for this plugin. Researchers interested in creating videos with other formats (e.g., AVI) can also use the image folders created by FaReT as input to the excellent open-source software ImageJ (<https://imagej.net>).

An advantage of creating dynamic face stimuli using FaReT rather than image morphs is that our toolkit uses linear interpolation in the shape and pose spaces, producing a more natural outcome than changes in an arbitrary morphing space, and takes advantage of MakeHuman’s algorithms aimed at simulating morphological features of the human face and its musculature. Recently, some researchers have questioned whether dynamic stimuli created using image morphs can capture the true dynamics of natural emotional expressions (Bernstein and Yovel, 2015; Roesch et al., 2011). The use of MakeHuman is a step in the right direction, especially if future applications precisely manipulate the speed of unfolding of emotional expressions to match those observed in natural face videos.

**Communicating with experiment software (e.g., PsychoPy) to render faces dynamically online** As indicated in the introduction, an important motivation behind developing FaReT was

**Table 3.** Emotion ratings provided by participants in Study 2

	N	Fear		Sadness		Anger		Happiness		Disgust		Surprise	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Happiness	33	1.29	0.90	1.18	0.59	1.09	0.45	<u>4.24</u>	1.87	1.17	0.59	1.44	0.99
Disgust	32	1.79	1.53	1.40	0.93	<b>4.96</b>	1.91	1.07	0.34	<u>4.58</u>	2.05	1.63	1.28

Each row represents a different expression model shown to the participants. Columns represent different expression ratings provided by the participants for the displayed face stimuli. For each identity that was morphed to a specific emotional expression, the participants provided ratings for six emotions on a scale from 1 to 7. For each expression model, the rating corresponding to the intended expression is underlined, and the highest-rated expression is in bold

to be able to freely manipulate and control face stimuli to such an extent that traditional psychophysical techniques would be available to face researchers. Among such techniques, some of the most useful require online generation of stimuli as a function of the participant's behavior, including adaptive procedures to estimate thresholds, psychophysical functions, or internal representations (Leek, 2001; Lu and Doshier, 2013; Shen, 2013; Treutwein, 1995; Watson and Pelli, 1983; Watson, 2017).

By implementing a plugin that uses the Python socket module, MakeHuman can respond to external requests to generate and render faces from external experiment control software like PsychoPy (Peirce, 2007, 2009). By generating stimuli online, it becomes possible to run adaptive psychophysics experiments. This tool allows researchers for the first time to easily program adaptive psychophysics experiments involving complex visual objects. However, because of the cumulative nature of rendering time, we recommend generating stimuli that use single images per trial for online experimental paradigms (especially if rendering with advanced lighting settings). For the time being, experiments that require stimuli with many frames (i.e., dynamic face stimuli) should be constrained to pre-generated sets of stimuli in most systems. The fast rendering required to generate dynamic stimuli online is possible, but it is likely to require both powerful hardware and additional programming (e.g., by parallelizing the task of rendering a sequence of images).

Instructions and examples on how to manipulate shape and expression parameters from Python code using the Socket Render plugin are available in the FaReT GitHub page: <https://github.com/fsotoc/FaReT#communicating-with-psycho-py-to-render-faces-online>. The plugin allows one to load MakeHuman identity and expression models, modify their parameters, and render the modified model online. It is also possible to create the models from scratch within Python, by simply generating dictionaries of parameter values. To use as a reference, we include tables with all the relevant parameters for the models of face shape and expression (see Appendices 1 and 2), together with the name given to the corresponding feature in MakeHuman's GUI, and the range of values that the parameter accepts. This will allow researchers to use the MakeHuman GUI to understand how different parameters change face shape and expression, and then use that knowledge to directly change the parameters from within their Python code.

### Installing and learning to use FaReT

The latest version of FaReT can be downloaded from the following link: <https://github.com/fsotoc/FaReT>. The page includes instructions on how to install and use the different

components of the toolkit. A folder is included with our database of MakeHuman models of face identity and expression.

## Study 1: database validation

We performed a validation study to determine whether dynamic stimuli created using FaReT are reliably perceived by human participants. If our synthesized identities and expressions are indeed perceived correctly, then we would expect that facial expressions would receive the highest ratings for the emotion that they actually represent and that the participants would rate the male identities as more masculine than the female identities. Both of these criteria have been previously used to validate computer-generated face stimuli (Roesch et al., 2011). In addition, we wanted to know to what extent the stimuli generated with FaReT are considered to be unusual by human participants. Ideally, computer-generated face stimuli would be comparable to naturalistic photos, but it seems clear that this level of realism cannot be achieved by currently available software (e.g., MakeHuman or FaceGen). However, it would be useful to quantitatively evaluate the level to which participants judge our stimuli to be unusual compared with naturalistic stimuli. With that goal in mind, our study was modeled after a previous validation study by Ma et al. (2015), who collected norming data for photographs of faces. We used their questions so that we could compare the responses about our computer-generated face images to the responses about their naturalistic face images.

## Method

**Participants** One hundred and twenty-three students from Florida International University (located in Miami, Florida) were recruited to participate in this study in exchange for credit. Thirty-eight participants tried to take the survey multiple times after being removed by catch questions (see description below) that assessed whether they were actually reading the questions. Six additional participants did not finish the survey and were also removed. The remaining 79 participants (71 female participants) finished the entire survey.

The participants' ages ranged from 18 to 45 years ( $M = 22.28$ ,  $SD = 4.82$ ). The participants' reported ethnicity distribution was as follows: 5 participants (6.33%) reported Asian, 15 reported Black or African American (18.99%), 33 (41.77%) reported White, 2 (2.53%) reported Native Hawaiian or Pacific Islander, 47 (59.49%) reported "Hispanic or Latino(a)", and 1 (1.27%) reported "Other".

**Stimuli** The face stimuli shown to participants were animated GIFs, created from renders of one identity model showing one

of the six expression pose models (i.e., happy, sad, afraid, surprised, angry, or disgusted). The animation involved a simple camera rotation, with the camera orbiting the viewpoint back and forth horizontally between  $-30$  and  $30$  degrees. The full camera motion consisted of 120 frames shown at 30 frames per second. The camera started from a frontal viewpoint, and the full camera movement loop lasted four seconds.

All 27 identity models were shown to participants: 12 male models, 12 female models, average male, average female, and global average. All possible combinations of identity and expression models were used, for a total of 162 animated stimuli. However, not every participant saw the same combination of identity and expression. Rather, each participant rated a subset of 23 face/expression combinations, which were selected randomly without replacement across participants. This sampling ensured that all face/expression combinations were shown once before the sampling process was started again.

**Procedure** The study was performed online through the Qualtrics survey program (<https://www.qualtrics.com/>). Participants were tasked with rating identities from our database by answering 13 questions per stimulus, in which they were asked to rate the extent to which the identities appeared threatening, attractive, baby-faced, trustworthy, masculine, fearful, sad, angry, happy, surprised, disgusted, and unusual on a scale from 1 to 7 (1 being the lowest and 7 being the highest). Participants also rated the identity's Euro/Afrocentricity on a scale from 1 to 100. As indicated earlier, the study included the same questions as a prior validation study by Ma et al. (2015). Responses to all questions are available to interested researchers (raw data can be downloaded from <https://osf.io/grp9d/>), but here we report analyses of only some of the questions, aimed at validating the stimuli generated by FaReT.

Each participant only had to rate 23 of our stimuli, which were intended to be pseudo-randomly assigned to ensure that the stimuli were all rated before reshown one. There were four catch trials inserted randomly among the other trials, with the goal of determining whether participants were paying attention to the rating task, rather than responding randomly without regard for instructions. In catch trials, participants were simply asked to select a specific numerical rating on the 1-to-7 scale (even though no dimension was specified). If participants failed to provide the specified rating, they were redirected to the end of the survey, and their data was not included in the results. The same two stimuli were always presented during catch trials: a surprised female model and a happy global average model.

## Results and discussion

One criterion used in the past to determine whether participants reliably perceive synthetic identities is that they

unambiguously recognize them as either male or female (Roesch et al., 2011). Seventy-nine participants provided masculinity ratings, and they rated the masculinity of male identities ( $M = 5.32$ ,  $SD = 1.09$ ) significantly higher than that of female identities ( $M = 3.53$ ,  $SD = 1.19$ );  $t(78) = 13.01$ ,  $p < .001$ , Cohen's  $d = 1.46$ . This implies that the participants were able to differentiate males from females in our database despite the lack of geometry-based cues (i.e., hair).

A second criterion used in prior research to validate computer-generated face stimuli is checking that participants rate the displayed expression most highly in its corresponding emotional category (e.g., faces expressing fear are given a relatively high fearful rating, faces expressing happiness are given a relatively high happiness rating, etc.; see Roesch et al., 2011). As shown in Table 1, this was true of our stimuli for the most part. When participants were shown fearful faces, they rated them higher in fear expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 438) = 19.74$ ,  $p < .001$ ,  $\eta_p^2 = .1839$ , and by a planned contrast comparing mean fear rating versus mean rating of all other emotions,  $F(1, 438) = 40.52$ ,  $p < .001$ ,  $\eta_p^2 = .0847$ . When participants were shown sad faces, they rated them higher in sadness expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 438) = 32.68$ ,  $p < .001$ ,  $\eta_p^2 = .2717$ , and by a planned contrast comparing mean sadness rating versus mean rating of all other emotions,  $F(1, 438) = 122.11$ ,  $p < .001$ ,  $\eta_p^2 = .2180$ . When participants were shown angry faces, they rated them higher in anger expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 426) = 39.37$ ,  $p < .001$ ,  $\eta_p^2 = .3160$ , and by a planned contrast comparing mean anger rating versus mean rating of all other emotions,  $F(1, 426) = 152.8$ ,  $p < .001$ ,  $\eta_p^2 = .2639$ . When participants were shown happy faces, they rated them higher in happiness expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 444) = 22.2$ ,  $p < .001$ ,  $\eta_p^2 = .2$ , and by a planned contrast comparing mean happiness rating versus mean rating of all other emotions,  $F(1, 444) = 110.4$ ,  $p < .001$ ,  $\eta_p^2 = .1991$ . When participants were shown disgusted faces, they rated them higher in anger expression than in any other expression, and second highest in disgust expression. The difference in ratings was significant according to a repeated-measures ANOVA,  $F(5, 438) = 35.39$ ,  $p < .001$ ,  $\eta_p^2 = .2877$ . A planned contrast showed that that disgusted faces were rated significantly higher in disgust expression than in all other expressions combined (i.e., on average),  $F(1, 438) = 14.98$ ,  $p < .001$ ,  $\eta_p^2 = .0331$ .

Finally, when participants were shown surprised faces, they rated them higher in surprise expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 438) = 35.8, p < .001, \eta_p^2 = .2901$ , and by a planned contrast comparing mean surprise rating versus mean rating of all other emotions,  $F(1, 438) = 127.4, p < .001, \eta_p^2 = .2253$ .

In sum, for the most part, the participants rated the displayed expression most highly in its corresponding emotional category. However, participants rated identities expressing disgust as showing a higher anger expression than a disgust expression (see Table 1). Given that our expression pose models were developed based on the photographs from a single actor from the KDEF database (Lundqvist et al., 1998), one possibility to consider is that the poor recognition of disgust in our study could stem from the ambiguity of this expression in the original photographs of that particular actor. This is highly unlikely, as norms have been published for the KDEF database (Goeleven et al., 2008), and people consistently classified the actor's expression as disgusted (hit rate above 90%), with the most common misclassification error being anger. We discuss other explanations for this specific result later.

The mean unusualness ratings for our identities were right in the middle of the scale ( $M = 4.09, SD = 1.06$ ), which was significantly higher than the ratings assigned to the photographs in the Ma et al. (2015) study ( $M = 2.56$ );  $t(78) = 12.82, p < .001$ , Cohen's  $d = 1.44$ . In that study, participants rated face photographs in the lower end of unusualness (i.e., well below the midpoint of 4 in the scale). Assuming that our participant population is similar to that used in the Ma et al. (2015) study (it might not be, as their demographics differ considerably), this implies that our identities are more unusual than the naturalistic photos, which could be due to the lack of hair, lighting/shading effects, etc. We reasoned that if lack of hair contributed to the relatively high unusualness ratings, then as men are more prone to balding, their unusualness ratings should be lower than those of women. We performed an exploratory test to compare the unusualness ratings of male ( $M = 4.20, SD = 1.12$ ) and female ( $M = 4.07, SD = 1.24$ ) identities, but the difference was nonsignificant;  $t(78) = 1.06, p = 0.29$ , Cohen's  $d = .12$ .

Table 2 shows the mean unusualness of the stimuli broken down by expression pose model. It can be seen that, on average, some models had higher unusualness than others. The two highest-rated models were happiness and disgust. As mentioned before, these models seemed problematic for other reasons as well. In the case of happiness, the model was manually modified after judging that its application to some identities could produce off-putting results. We might have been unable to solve this issue completely. In the case of disgust, this model produced much higher ratings of anger than disgust

from participants (see Table 1). Although confusion between disgust and anger is commonly found in the literature, even with naturalistic photographic stimuli, it is problematic that the mean anger rating for these stimuli is about 1.5 points higher than the mean disgust rating. To address both issues, we built two new expression pose models and validated them in a new study that is presented in the following section.

## Study 2: validation of alternative models of happiness and disgust

As mentioned above, the results of Study 1 prompted us to develop two new expression pose models of happiness and disgust. For happiness, research has shown that open-mouth smiles in three-dimensional models are judged as more authentic than closed-mouth smiles (Korb et al., 2014), and thus we built a model showing an open-mouth smile. For disgust, we reasoned that we needed to build a model showing a feature that is only found in disgust and not in anger. Rozin et al. (1994) showed that there are a variety of ways to express disgust depending on the presence or absence of upper lip retraction, nose wrinkle, and gape and tongue extension. Each of these disgust expressions carry a different meaning, with the upper lip retraction carrying information about disgust in social situations (e.g., breaking of moral rules), the nose wrinkle communicating a bad smell (and to a lesser extent a bad taste), and the gape and tongue extension communicating oral irritation and the need to expel contaminated food from the mouth (Rozin and Fallon, 1987; Rozin et al., 1994). Upper lip retraction is common in expression of anger and contempt, which are important in social communication. Probably due to constraints in the structure of the face (Rozin et al., 1994), a nose wrinkle tends to be accompanied by a general contraction of face features thought to be related to sensory rejection (Susskind et al., 2008). Such contraction can also be found in expression of anger and contempt. On the other hand, gape and tongue extension is exclusively found in disgust and not in anger or contempt. Thus, we chose to include these features in the new disgust model, while noting that the new model involves a different form of disgust (related to contaminated food) than the model presented earlier (related to social communication and sensory rejection).

Renders of the two new models for happiness and disgust are shown in Fig. 10. The study used the exact same procedures as Study 1, but with stimuli generated using the new expression models.

## Method

**Participants** Thirty-three people participated in this study. The participants' ages ranged from 18 to 43 years ( $M = 23.5$ ,

$SD=5.19$ ), and their reported ethnicity distribution was as follows: 2 participants (6.06%) reported Asian, 4 reported Black or African American (12.12%), 8 (24.24%) reported White, 0 (0%) reported Native Hawaiian or Pacific Islander, 25 (75.75%) reported “Hispanic or Latino(a)”, and 3 (9.09%) reported “Other”. Other features of the sample were as described for Study 1.

**Stimuli** Stimuli were generated as described for Study 1, but using the two new models for happiness and disgust displayed in Fig. 10.

**Procedures** Procedures were as described for Study 1, with the exception that here every participant rated every stimulus (i.e., combination of identity model and the two new expression pose models).

## Results and discussion

As in Study 1, the 31 participants who provided masculinity ratings rated the masculinity of male identities ( $M=4.93$ ,  $SD=1.66$ ) significantly higher than that of female identities ( $M=3.05$ ,  $SD=1.57$ );  $t(30)=7.61$ ,  $p<.001$ , Cohen’s  $d=2.79$ . As shown in Table 3, when participants were shown happy faces with open-mouth smiles, they rated them higher in happiness expression than in any other expression. This was confirmed by a repeated-measures ANOVA,  $F(5, 186)=7.59$ ,  $p<.001$ ,  $\eta_p^2=.1694$ , and by a planned contrast comparing mean happiness rating versus mean rating of all other emotions,  $F(1, 186)=36.41$ ,  $p<.001$ ,  $\eta_p^2=.1637$ . When participants were shown disgusted faces with an open mouth and the tongue sticking out, they rated them highest in anger expression, just as in the previous study. However, unlike Study 1, anger ratings were much lower and close to disgust ratings. The repeated-measures ANOVA showed a significant difference in ratings of different emotions,  $F(5, 180)=21.60$ ,  $p<.001$ ,  $\eta_p^2=.3750$ , and the planned contrast comparing mean disgust rating versus mean rating of all other emotions was also significant,  $F(1, 180)=39.95$ ,  $p<.001$ ,  $\eta_p^2=.1816$ . Thus, an improvement was observed in this study regarding ratings of the disgust expression model. However, we were unable to eliminate confusion between anger and disgust. As indicated earlier, this is a common finding in the literature using naturalistic stimuli (see General Discussion below).

The mean unusualness ratings for our stimuli was on average lower than that of the previous study ( $M=3.36$ ,  $SD=1.33$ ), but still significantly higher than the ratings assigned to the photographs in the Ma et al. (2015) study ( $M=2.56$ );  $t(33)=3.48$ ,  $p<.01$ , Cohen’s  $d=1.21$ . As shown in Table 2, both models showed lower unusualness than their Study 1

counterparts, although the happiness model showed a larger drop in rating.

Overall, from these results it does seem advisable to use the open-mouth models depicted in Fig. 10 rather than the originally developed models for happiness and disgust depicted in Fig. 3, although further improvements are likely possible toward a completely unambiguous model of disgust.

## General discussion

Here, we presented FaReT (Face Research Toolkit), a free and open-source toolkit of 3D models and software to study face perception. FaReT allows face perception researchers to easily manipulate 3D models of faces in ways that are common in the literature (morphing, averaging, standardization, etc.), but with improved speed, control, and interpretability of such manipulations. FaReT is built on top of MakeHuman, taking advantage of this software’s intuitive shape and pose spaces, where transformations are natural and interpretable. Our toolkit also easily produces dynamic stimuli that are more natural and interpretable than those produced through concatenation of image morphs, but also more controlled and manipulable than natural videos. The toolkit includes applications to enable easy control of relevant face features and standardization of irrelevant features, where “relevant” and “irrelevant” can be defined by research goals. Finally, the toolbox allows easy implementation of adaptive psychophysical techniques in the study of face perception, which traditionally have been available only with very simple, low-level stimuli.

FaReT also includes a database of 27 face identity models and six expression pose models. We performed a validation study in which the models in our database were used to create dynamic facial stimuli that participants rated in a number of perceptual features. We found that raters typically assigned the highest scores to emotions that aligned with the artificial stimuli’s corresponding facial expression. The exception was our pose model for disgust, which was rated highest in anger. Throughout the face research literature, anger and disgust are often rated similarly or confused (Thorstenson et al., 2019), including in previous studies involving three-dimensional face models (see Roesch et al., 2011, Fig. 5). Studies using adaptation suggest that they may be encoded through overlapping channels (Skinner and Benton, 2010), meaning that the artificial expressions of these emotions were portrayed well enough to capture that similarity. In the same vein, the expressions of fear and surprise also received similar ratings of emotionality.

Some of the functionality provided by FaReT is also included in the FACSHuman plugin (Gilbert et al., 2018; <https://github.com/montybot/FACSHuman>). For example, both provide the ability to create image sequences and videos through linear interpolation in model space. However, each package has its own strengths. In FACSHuman, expression modeling is encoded using the FACS system, which is familiar to many researchers in face perception. Some existing face image data sets are already FACS-rated (e.g., Lucey et al., 2010; Mavadati et al., 2013), which would allow for easy development of new valid models of emotional expression using FACSHuman.

On the other hand, FaReT was designed with the broader goal of helping face perception researchers produce controlled stimuli in psychophysical experiments, using tasks and designs already common in the current literature, as well as those that have been previously available only for research in low-level vision. Given that goal, FaReT allows manipulation of not only expression parameters, but also identity parameters, and includes linear extrapolation in addition to interpolation, allowing researchers to create anti-models and caricature models. This, plus the ability to easily generate average models and standardize large sets of models, allows FaReT to perform all stimulus manipulations that researchers in face perception currently achieve using morphing software. Additionally, FaReT includes a plugin that enables interactive communication with MakeHuman from Python, which opens the possibility to design adaptive psychophysics experiments using software such as Psychopy. In combination with flexible adaptive procedures such as QUEST+ (Watson, 2017; <https://github.com/hoechenberger/questplus/tree/master/questplus>), researchers have access to fast estimation of parameters of any parametrically defined psychophysical function, such as psychometric curves, threshold vs. stimulus parameter functions (i.e., similar to contrast sensitivity functions), transducer functions, threshold versus external noise functions (both similar to those estimated using contrast in low-level vision), and parameters of observer models (see Lu and Doshier, 2013).

Despite these successes, one limitation of FaReT highlighted by our results is that the artificial stimuli were given ratings of unusualness that are higher than those previously provided for static photographs of human faces (Ma et al., 2015). At this point, it is not clear what aspects of our stimuli or procedure might have produced the difference, as the current study and that of Ma et al. (2015) differed in several ways besides the synthetic versus natural faces. However, the assumption that the result is mostly due to the synthetic nature of our stimuli seems reasonable, and in that case a clear goal for future releases of FaReT is enhancing the quality of the assets used to generate the faces (such as skin textures), deriving

parameters for faces and expressions from infrared depth-based images, or altering the type of camera used to create renders (from MakeHuman's hard-coded orthographic camera to a perspective camera).

A second limitation of the current version of FaReT is that its model database has a limited number of identities and expressions. As a result, the average identities and expressions are likely to be biased estimates of the true averages (Cook et al., 2011). Fortunately, this database can be expanded without limit by the scientific community. Indeed, the larger community of MakeHuman users has already provided an extensive and growing database of identity (<http://www.makehumancommunity.org/models.html>) and expression (<http://www.makehumancommunity.org/expressions.html>) models, as well as other useful assets and plugins (see [http://www.makehumancommunity.org/content/user\\_contributed\\_assets.html](http://www.makehumancommunity.org/content/user_contributed_assets.html)). In addition, the fact that models in MakeHuman are described in a multidimensional space of shape parameters offers the possibility of taking a much more useful approach to this problem: fully characterizing the distribution of realistic models of identity and expression in such space, rather than any individual model, and sampling novel identities and expressions from that distribution. This would allow not only a larger database, but actually the generation of an infinite number of novel realistic models. We expect this to be the next step of development in FaReT.

We hope that FaReT will help face perception researchers in solving some of the methodological problems that they currently face, as well as opening new avenues of research, all while maintaining an open-source status. We want to allow users to precisely modify 3D facial features through MakeHuman's sensibly named parameters, without depriving them of the important morphing procedures that had originally arisen from image manipulation software. This level of quantitative control is imperative to successful psychometric research. Our goal is to provide vision scientists with malleable high-level visual stimuli while maintaining the benefits of control that low-level vision scientists enjoy. We hope that this will help the scientific community to answer bigger questions without paying the unnecessary costs (both financial and scientific) imposed by copyrighted software.

**Acknowledgements** Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R21MH112013 to Fabian A. Soto. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. *S.D.G.*

**Open Practices Statement** The data and materials for the validation studies reported here are available at <https://osf.io/grp9d/>. These studies were not preregistered. The FaReT toolbox described here and used to generate stimuli for the validation study can be found at <https://github.com/fsotoc/FaReT>, which is linked in the OSF study page as well.

## Appendix 1

**Table 4** Parameters of the Face Identity Model

GUI Category	GUI Label	Parameter Key	Range
Head shape	Age	head/head-age-decr incr	(-1,1)
Head shape	Head fat	head/head-fat-decr incr	(-1,1)
Head shape	Angle	head/head-angle-in out	(-1,1)
Head shape	Oval	head/head-oval	(0,1)
Head shape	Round	head/head-round	(0,1)
Head shape	Rectangular	head/head-rectangular	(0,1)
Head shape	Square	head/head-square	(0,1)
Head shape	Triangular	head/head-triangular	(0,1)
Head shape	Inverted triangular	head/head-invertedtriangular	(0,1)
Head shape	Diamond	head/head-diamond	(0,1)
Head shape	Scale depth of parietal side	head/head-back-scale-depth-decr incr	(-1,1)
Head size	Scale depth	head/head-scale-depth-decr incr	(-1,1)
Head size	Scale horizontally	head/head-scale-horiz-decr incr	(-1,1)
Head size	Scale vertically	head/head-scale-vert-decr incr	(-1,1)
Head size	Move horizontally	head/head-trans-in out	(-1,1)
Head size	Move vertically	head/head-trans-down up	(-1,1)
Head size	Move depth	head/head-trans-backward forward	(-1,1)
Forehead	Forehead bulge	forehead/forehead-trans-backward forward	(-1,1)
Forehead	Scale vertically	forehead/forehead-scale-vert-decr incr	(-1,1)
Forehead	Cranic shape	forehead/forehead-nubian-decr incr	(-1,1)
Forehead	Temple bulge	forehead/forehead-temple-decr incr	(-1,1)
Eyebrows	Eyebrows bulge	eyebrows/eyebrows-trans-backward forward	(-1,1)
Eyebrows	Eyebrows angle	eyebrows/eyebrows-angle-down up	(-1,1)
Eyebrows	Move vert	eyebrows/eyebrows-trans-down up	(-1,1)
Neck	Double neck	neck/neck-double-decr incr	(-1,1)
Neck	Scale depth	neck/neck-scale-depth-decr incr	(-1,1)
Neck	Scale horizontally	neck/neck-scale-horiz-decr incr	(-1,1)
Neck	Scale vertically	neck/neck-scale-vert-decr incr	(-1,1)
Neck	Move horizontally	neck/neck-trans-in out	(-1,1)
Neck	Move vertically	neck/neck-trans-down up	(-1,1)
Neck	Move depth	neck/neck-trans-backward forward	(-1,1)
Neck	Scale depth of nape	neck/neck-back-scale-depth-decr incr	(-1,1)
Right eye	Eye bag volume	eyes/r-eye-bag-decr incr	(-1,1)
Right eye	Eye bag distorsion	eyes/r-eye-bag-in out	(-1,1)
Right eye	Eye bag height	eyes/r-eye-bag-height-decr incr	(-1,1)
Right eye	Eyefold angle	eyes/r-eye-eyefold-angle-down up	(-1,1)
Right eye	Eye epicanthus	eyes/r-eye-epicanthus-in out	(-1,1)
Right eye	Eyefold volume	eyes/r-eye-eyefold-concave convex	(-1,1)
Right eye	Eyefold position	eyes/r-eye-eyefold-down up	(-1,1)
Right eye	Scale height 1	eyes/r-eye-height1-decr incr	(-1,1)
Right eye	Scale height 2	eyes/r-eye-height2-decr incr	(-1,1)
Right eye	Scale height 3	eyes/r-eye-height3-decr incr	(-1,1)
Right eye	Move outer corner horiz.	eyes/r-eye-push1-in out	(-1,1)
Right eye	Move inner corner horiz.	eyes/r-eye-push2-in out	(-1,1)
Right eye	Move eye horizontally	eyes/r-eye-trans-in out	(-1,1)
Right eye	Move eye vertically	eyes/r-eye-trans-down up	(-1,1)
Right eye	Scale eye	eyes/r-eye-scale-decr incr	(-1,1)
Right eye	Move outer corner vert.	eyes/r-eye-comer1-down up	(-1,1)
Right eye	Move inner corner vert.	eyes/r-eye-comer2-down up	(-1,1)
Left eye	Eye bag volume	eyes/l-eye-bag-decr incr	(-1,1)
Left eye	Eye bag distorsion	eyes/l-eye-bag-in out	(-1,1)
Left eye	Eye bag height	eyes/l-eye-bag-height-decr incr	(-1,1)
Left eye	Eyefold angle	eyes/l-eye-eyefold-angle-down up	(-1,1)
Left eye	Eye epicanthus	eyes/l-eye-epicanthus-in out	(-1,1)
Left eye	Eyefold volume	eyes/l-eye-eyefold-concave convex	(-1,1)
Left eye	Eyefold position	eyes/l-eye-eyefold-down up	(-1,1)
Left eye	Scale height 1	eyes/l-eye-height1-decr incr	(-1,1)
Left eye	Scale height 2	eyes/l-eye-height2-decr incr	(-1,1)
Left eye	Scale height 3	eyes/l-eye-height3-decr incr	(-1,1)
Left eye	Move outer corner horiz.	eyes/l-eye-push1-in out	(-1,1)
Left eye	Move inner corner horiz.	eyes/l-eye-push2-in out	(-1,1)

**Table 4** (continued)

GUI Category	GUI Label	Parameter Key	Range
Left eye	Move eye horizontally	eyes/l-eye-trans-in out	(-1,1)
Left eye	Move eye vertically	eyes/l-eye-trans-down up	(-1,1)
Left eye	Scale eye	eyes/l-eye-scale-decr incr	(-1,1)
Left eye	Move outer corner vert.	eyes/l-eye-corer1-down up	(-1,1)
Left eye	Move inner corner vert.	eyes/l-eye-corer2-down up	(-1,1)
Nose size	Move vertically	nose/nose-trans-down up	(-1,1)
Nose size	Move depth	nose/nose-trans-backward forward	(-1,1)
Nose size	Move horizontally	nose/nose-trans-in out	(-1,1)
Nose size	Scale vertically	nose/nose-scale-vert-decr incr	(-1,1)
Nose size	Scale horizontally	nose/nose-scale-horiz-decr incr	(-1,1)
Nose size	Scale depth	nose/nose-scale-depth-decr incr	(-1,1)
Nose size details	Scale nostrils	width nose/nose-nostrils-width-decr incr	(-1,1)
Nose size details	Scale tip width	nose/nose-point-width-decr incr	(-1,1)
Nose size details	Move base vert.	nose/nose-base-down up	(-1,1)
Nose size details	Scale width 1	nose/nose-width1-decr incr	(-1,1)
Nose size details	Scale width 2	nose/nose-width2-decr incr	(-1,1)
Nose size details	Scale width 3	nose/nose-width3-decr incr	(-1,1)
Nose features	Compression	nose/nose-compression-compress uncompress	(-1,1)
Nose features	Curve	nose/nose-curve-concave convex	(-1,1)
Nose features	Greek	nose/nose-greek-decr incr	(-1,1)
Nose features	Hump	nose/nose-hump-decr incr	(-1,1)
Nose features	Volume	nose/nose-volume-decr incr	(-1,1)
Nose features	Nostrils	nose/nose-nostrils-angle-down up	(-1,1)
Nose features	Move tip vertically	nose/nose-point-down up	(-1,1)
Nose features	Septum Angle	nose/nose-septumangle-decr incr	(-1,1)
Nose features	Scale nostrils flaring	nose/nose-flaring-decr incr	(-1,1)
Mouth size	Scale horizontally	mouth/mouth-scale-horiz-decr incr	(-1,1)
Mouth size	Scale vertically	mouth/mouth-scale-vert-decr incr	(-1,1)
Mouth size	Scale depth	mouth/mouth-scale-depth-decr incr	(-1,1)
Mouth size	Move horizontally	mouth/mouth-trans-in out	(-1,1)
Mouth size	Move vertically	mouth/mouth-trans-down up	(-1,1)
Mouth size	Move depth	mouth/mouth-trans-backward forward	(-1,1)
Mouth size details	Scale lowerlip height	mouth/mouth-lowerlip-height-decr incr	(-1,1)
Mouth size details	Scale lowerlip width	mouth/mouth-lowerlip-width-decr incr	(-1,1)
Mouth size details	Scale upperlip height	mouth/mouth-upperlip-height-decr incr	(-1,1)
Mouth size details	Scale upperlip width	mouth/mouth-upperlip-width-decr incr	(-1,1)
Mouth size details	Cupid's bow width	mouth/mouth-cupidsbow-width-decr incr	(-1,1)
Mouth features	Dimples	mouth/mouth-dimples-in out	(-1,1)
Mouth features	Laugh-lines	mouth/mouth-laugh-lines-in out	(-1,1)
Mouth features	Lowerlip curved shape	mouth/mouth-lowerlip-ext-down up	(-1,1)
Mouth features	Move corners vert.	mouth/mouth-angles-down up	(-1,1)
Mouth features	Scale middle lowerlip	mouth/mouth-lowerlip-middle-down up	(-1,1)
Mouth features	Scale lowerlip volume	mouth/mouth-lowerlip-volume-decr incr	(-1,1)
Mouth features	Scale philtrum volume	mouth/mouth-philtrum-volume-decr incr	(-1,1)
Mouth features	Scale upperlip volume	mouth/mouth-upperlip-volume-decr incr	(-1,1)
Mouth features	Upperlip curved shape	mouth/mouth-upperlip-ext-down up	(-1,1)
Mouth features	Scale middle upperlip	mouth/mouth-upperlip-middle-down up	(-1,1)
Mouth features	Cupid's bow shape	mouth/mouth-cupidsbow-decr incr	(-1,1)
Right ear	Move depth	ears/r-ear-trans-backward forward	(-1,1)
Right ear	Scale ear	ears/r-ear-scale-decr incr	(-1,1)
Right ear	Move vertically	ears/r-ear-trans-down up	(-1,1)
Right ear	Scale height	ears/r-ear-scale-vert-decr incr	(-1,1)
Right ear	Scale lobe	ears/r-ear-lobe-decr incr	(-1,1)
Right ear	Ear shape (pointed-triangle)	ears/r-ear-shape-pointed triangle	(-1,1)
Right ear	Ear rotation	ears/r-ear-rot-backward forward	(-1,1)
Right ear	Ear shape (squared-round)	ears/r-ear-shape-square round	(-1,1)
Right ear	Scale width	ears/r-ear-scale-depth-decr incr	(-1,1)
Right ear	Ear wing-shaped	ears/r-ear-wing-decr incr	(-1,1)
Right ear	Ear flapped	ears/r-ear-flap-decr incr	(-1,1)
Left ear	Move depth	ears/l-ear-trans-backward forward	(-1,1)
Left ear	Scale ear	ears/l-ear-scale-decr incr	(-1,1)
Left ear	Move vertically	ears/l-ear-trans-down up	(-1,1)
Left ear	Scale height	ears/l-ear-scale-vert-decr incr	(-1,1)
Left ear	Scale lobe	ears/l-ear-lobe-decr incr	(-1,1)
Left ear	Ear shape (pointed-triangle)	ears/l-ear-shape-pointed triangle	(-1,1)



**Table 4** (continued)

GUI Category	GUI Label	Parameter Key	Range
Left ear	Ear rotation	ears/l-ear-rot-backward forward	(-1,1)
Left ear	Ear shape (squared-round)	ears/l-ear-shape-square round	(-1,1)
Left ear	Scale width	ears/l-ear-scale-depth-decr incr	(-1,1)
Left ear	Ear wing-shaped	ears/l-ear-wing-decr incr	(-1,1)
Left ear	Ear flapped	ears/l-ear-flap-decr incr	(-1,1)
Chin/jaw	Tone of side chin	chin/chin-jaw-drop-decr incr	(-1,1)
Chin/jaw	Cleft chin	chin/chin-cleft-decr incr	(-1,1)
Chin/jaw	Scale chin prominence	chin/chin-prominent-decr incr	(-1,1)
Chin/jaw	Scale chin width	chin/chin-width-decr incr	(-1,1)
Chin/jaw	Scale chin height	chin/chin-height-decr incr	(-1,1)
Chin/jaw	Scale chin angular	chin/chin-bones-decr incr	(-1,1)
Chin/jaw	Scale chin prognathism	chin/chin-prognathism-decr incr	(-1,1)
Left cheek	Cheek outer volume	cheek/l-cheek-volume-decr incr	(-1,1)
Left cheek	Scale cheek prominence	cheek/l-cheek-bones-decr incr	(-1,1)
Left cheek	Cheek inner volume	cheek/l-cheek-inner-decr incr	(-1,1)
Left cheek	Move vertically	cheek/l-cheek-trans-down up	(-1,1)
Right cheek	Cheek outer volume	cheek/r-cheek-volume-decr incr	(-1,1)
Right cheek	Scale cheek prominence	cheek/r-cheek-bones-decr incr	(-1,1)
Right cheek	Cheek inner volume	cheek/r-cheek-inner-decr incr	(-1,1)
Right cheek	Move vertically	cheek/r-cheek-trans-down up	(-1,1)
Main/Macro	Gender	macrodetails/Gender	(0,1)
Main/Macro	Age	macrodetails/Age	(0,1)
Main/Macro	Muscle	macrodetails-universal/Muscle	(0,1)
Main/Macro	African	macrodetails/African	(0,1)
Main/Macro	Asian	macrodetails/Asian	(0,1)
Main/Macro	Caucasian	macrodetails/Caucasian	(0,1)

## Appendix 2

**Table 5** Parameters of the Expression Pose Model

GUI Label / Parameter Key	Range
CheeksPump	(0,1)
CheeksSuck	(0,1)
ChinDown	(0,1)
ChinForward	(0,1)
ChinLeft	(0,1)
ChinRight	(0,1)
JawDrop	(0,1)
JawDropStretched	(0,1)
LeftBrowDown	(0,1)
LeftCheekUp	(0,1)
LeftEyeDown	(0,1)
LeftEyeUp	(0,1)
LeftEyeturnLeft	(0,1)
LeftEyeturnRight	(0,1)
LeftInnerBrowUp	(0,1)
LeftLowerLidUp	(0,1)
LeftOuterBrow	(0,1)
Up LeftUpperLidClosed	(0,1)
LeftUpperLidOpen	(0,1)
LipsKiss	(0,1)
MouthLeftPlatysma	(0,1)
MouthLeftPullDown	(0,1)
MouthLeftPullSide	(0,1)
MouthLeftPullUp	(0,1)
MouthMoveLeft	(0,1)
MouthMoveRight	(0,1)
MouthRightPlatysma	(0,1)
MouthRightPullDown	(0,1)
MouthRightPullSide	(0,1)
MouthRightPullUp	(0,1)
NasolabialDeepener	(0,1)
NoseWrinkler	(0,1)
Rest	(0,1)
RightBrowDown	(0,1)
RightCheekUp	(0,1)
RightEyeDown	(0,1)
RightEyeUp	(0,1)
RightEyeturnLeft	(0,1)
RightEyeturnRight	(0,1)
RightInnerBrowUp	(0,1)
RightLowerLidUp	(0,1)
RightOuterBrowUp	(0,1)
RightUpperLidClosed	(0,1)
RightUpperLidOpen	(0,1)
TongueDown	(0,1)
TongueLeft	(0,1)
TongueOut	(0,1)
TonguePointDown	(0,1)
TonguePointUp	(0,1)
TongueRight	(0,1)
TongueUp	(0,1)
TongueUshape	(0,1)
UpperLipBackward	(0,1)
UpperLipForward	(0,1)
UpperLipStretched	(0,1)
UpperLipUp	(0,1)
lowerLipBackward	(0,1)
lowerLipDown	(0,1)
lowerLipForward	(0,1)
lowerLipUp	(0,1)

## References

- Aguado, L., Garcia-Gutierrez, A., and Serrano-Pedraza, I. (2009). Symmetrical interaction of sex and expression in face classification tasks. *Attention, Perception & Psychophysics*, 71(1): 9.
- Anzellotti, S. and Caramazza, A. (2014). The neural mechanisms for the recognition of face identity in humans. *Frontiers in Psychology*, 5: 672.
- Bastioni, M., Re, S., and Misra, S. (2008). Ideas and methods for modeling 3d human figures: the principal algorithms used by MakeHuman and their implementation in a new approach to parametric modeling. In *Proceedings of the 1st Bangalore Annual Compute Conference*, pages 10:1–10:6, New York. ACM.
- Bayet, L., Pascalis, O., Quinn, P. C., Lee, K., Gentaz, E., and Tanaka, J. W. (2015). Angry facial expressions bias gender categorization in children and adults: behavioral and computational evidence. *Frontiers in Psychology*, 6:346.
- Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., and Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, 92(2):179–190.
- Bernstein, M. and Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, 55:536–546.
- Burton, N., Jeffery, L., Calder, A. J., and Rhodes, G. (2015). How is facial expression coded? *Journal of Vision*, 15(1):1–1.
- Byatt, G. and Rhodes, G. (1998). Recognition of own-race and other-race caricatures: implications for models of face recognition. *Vision Research*, 38(15):2455–2468.
- Calder, A. J., Burton, A. M., Miller, P., Young, A. W., and Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research*, 41(9):1179–1208.
- Ceipidor, U. B., Medaglia, C. M., Passacantilli, E., Fabri, S., Perrone, A., and Bastioni, M. (2008). Design of a GUI for the facial expressions creation in the 3d software “Make Human”-Demo. *Interaction Design and Architecture*, (5-6):121–122.
- Cook, R., Matei, M., and Johnston, A. (2011). Exploring expression space: Adaptation to orthogonal and anti-expressions. *Journal of Vision*, 11(4):1–9.
- Dailey, M., Cottrell, G. W., and Reilly, J. (2001). California facial expressions, CAFE.
- Duchaine, B. and Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, 1(1):393–416.
- Ebner, N. C., Riediger, M., and Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1):351–362.
- Ekman, P. (1999). Basic emotions. In Dalglish, T. and Power, M. J., editors, *Handbook of cognition and emotion*, pages 45–60. John Wiley & Sons Ltd, New York, NY, US.
- Ekman, P. and Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Unmasking the face: A guide to recognizing emotions from facial clues. Prentice-Hall, Oxford, England.
- Gilbert, M., Demarchi, S., and Urdapilleta, I. (2018). FACSHuman a software to create experimental material by modeling 3d facial expression. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents - IVA '18*, pages 333–334, Sydney, NSW, Australia. ACM Press.
- Goeleven, E., De Raedt, R., Leyman, L., and Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. *Cognition and Emotion*, 22(6):1094–1118.
- Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research*, 41(17): 2261–2271.

- Ho, P. K., Woods, A., and Newell, F. N. (2018). Temporal shifts in eye gaze and facial expressions independently contribute to the perceived attractiveness of unfamiliar faces. *Visual Cognition*, 26(10): 831–852.
- Kingdom, F. A. A. and Prins, N. (2016). *Psychophysics: A Practical Introduction*. Academic Press, Amsterdam, 2 edition.
- Korb, S., With, S., Niedenthal, P., Kaiser, S., and Grandjean, D. (2014). The perception and mimicry of facial movements predict judgments of smile authenticity. *PLOS ONE*, 9(6):e99194.
- Lamer, S. A., Weisbuch, M., and Sweeny, T. D. (2017). Spatial cues influence the visual perception of gender. *Journal of Experimental Psychology: General*, 146(9):1366–1371.
- Lander, K. and Butcher, N. (2015). Independence of face identity and expression processing: exploring the role of motion. *Frontiers in Psychology*, 6:255.
- Lee, K., Byatt, G., and Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: Testing the face-space framework. *Psychological Science*, 11(5):379–385.
- Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63(8):1279.
- Leopold, D. A., O'Toole, A. J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat Neurosci*, 4(1):89–94.
- Lu, Z. L. and Doshier, B. (2013). *Visual Psychophysics: From Laboratory to Theory*. MIT Press.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., and Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 94–101.
- Lundqvist, D., Flykt, A., and Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*, pages ISBN 91-630-7164-9.
- Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4):1122–1135.
- Macke, J. H. and Wichmann, F. A. (2010). Estimating predictive stimulus features from psychophysical data: The decision image technique applied to human faces. *Journal of Vision*, 10(5):22.
- Mangini, M. C. and Biederman, I. (2004). Making the ineffable explicit: estimating the information employed for face classifications. *Cognitive Science*, 28(2):209–226.
- Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P., and Cohn, J. F. (2013). DISFA: A spontaneous facial action intensity database. *IEEE Transactions on A ctive Computing*, 4(2):151–160.
- O'Toole, A. J., Abdi, H., Deffenbacher, K. A., and Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America*, 10(3):405–411.
- Oosterhof, N. N. and Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32):11087–11092.
- Pandzic, I. S. and Forchheimer, R., editors (2002). *MPEG-4 facial animation: The standard, implementation and applications*. Wiley, Hoboken, NJ, 1 edition.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2):8–13.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2:10.
- Rhodes, G. (2017). Adaptive coding and face recognition. *Current Directions in Psychological Science*, 26(3):218–224.
- Rhodes, G. and Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, 46(18):2977–2987.
- Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., and Scherer, K. R. (2011). FACSGen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *Journal of Nonverbal Behavior*, 35(1):1–16.
- Rozin, P. and Fallon, A. E. (1987). A perspective on disgust. *Psychological Review*, 94(1):23–41.
- Rozin, P., Lowery, L., and Ebert, R. (1994). Varieties of disgust faces and the structure of disgust. *Journal of Personality and Social Psychology*, 66(5):870–881.
- Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, 32(9):1093–1107.
- Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception*, 38(8):1211–1219.
- Schyns, P. G., Bonnar, L., and Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13(5):402–409.
- Shen, Y. (2013). Comparing adaptive procedures for estimating the psychometric function for an auditory gap detection task. *Attention, Perception, & Psychophysics*, 75(4):771–780.
- Skinner, A. L. and Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science*, 21(9):1248–1253.
- Soto, F. A. (2019). Categorization training changes the visual representation of face identity. *Attention, Perception, & Psychophysics*, 81(5): 1220–1227.
- Soto, F. A. and Ashby, F. G. (2015). Categorization training increases the perceptual separability of novel dimensions. *Cognition*, 139:105–129.
- Soto, F. A. and Ashby, F. G. (2019). Novel representations that support rule-based categorization are acquired on-the-fly during category learning. *Psychological Research*, 83(3):544–566.
- Steyvers, M. (1999). Morphing techniques for manipulating face images. *Behavior Research Methods*, 31(2):359–369.
- Strohming, N., Gray, K., Chituc, V., Heffner, J., Schein, C., and Heagins, T. B. (2016). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods*, 48(3):1197–1204.
- Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., and Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience*, 11(7):843–850.
- Thorstenson, C. A., Pazda, A. D., Young, S. G., and Elliot, A. J. (2019). Face color facilitates the disambiguation of confusing emotion expressions: Toward a social functional account of face color in emotion communication. *Emotion*, 19(5):799–807.
- Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35(17):2503–2522.
- Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86.
- Uddenberg, S. and Scholl, B. J. (2018). Teleface: Serial reproduction of faces reveals a whiteward bias in race memory. *Journal of Experimental Psychology: General*, 147(10):1466–1487.
- Watson, A. and Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2):113–120.
- Watson, A. B. (2017). QUEST+: A general multidimensional Bayesian adaptive psychometric method. *Journal of Vision*, 17(3):10–10.
- Webster, M. A. and MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571):1702–1725.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.