Entering the Design Space of Digital Portraiture:
A Case Study in Avatar Creation Tools

Jasmine T. Otto, Angus G. Forbes

University of California, Santa Cruz, 1156 High St, Santa Cruz, CA 95064, USA

Abstract
We describe creativity support tools for digital portraiture in terms of their design spaces, which conveniently possess both a high-dimensional linear space of representations, and a highly expressive nonlinear space of referents. Using the casual creators tropes of parameterization and paper dolls, we investigate how distinct ‘face spaces’ are made playable by these tools and toys. We consider in depth The Sims 3’s Create-A-Sim mode, a 3D avatar creator whose expressive range is further expanded by a modding community through intentional and highly technical interventions. Finally, we connect these skills to practices using professional creativity support tools for digital portraiture. We argue that the skillful practice of creating beautiful and evocative avatars is encoded simultaneously in software tools and in artistic communities, enabling a form of conversation between the portrait artist and their medium.

Keywords
avatar creation, identity play, design space, co-creation, skill development, the sims, modding

1. Introduction
Art students undertake years of life drawing to learn how to capture human likenesses in graphite and clay. Yet videogames which allow players to create their own character must somehow facilitate avatar creation by novice users, without making a hash of their first attempt with an unfamiliar system. Therefore, videogames with avatar customization invariably supply a combination of ‘known good’ starting avatars, and possibly an additional system of procedural modification [1].

These systems enable the exploration of a certain ‘face space’ - far being from the only one that exists, but certainly compatible with e.g. facial animations specific to the game, - which may not contain only good avatars. Indeed, more permissive editors tend to be capable of producing a wide variety of ‘monstrous’ faces. The humorous exploration and celebration of these possibilities formed the premise of the ‘Monster Factory’ Let’s Play series [2]. This paper will explore an argument that the existence and accessibility of these distressing and typically unintended results is indeed necessary and co-constitutive of the value in a creativity support system for digital software.

In order to make this claim, we need to describe both the technical and expressive dimensions of portraiture. We will focus on the case of 3D polygon-based avatars, because this is the current
technological standard for roleplaying games produced by major studios, and in order to use The Sims 3 (TS3) as our case study. While 2D avatars present a compelling example in the domains of videogaming, tabletop roleplaying, and social media, we leave a detailed exploration to the craft literature of the respective software, and therefore leave their technical description aside.

Faces are highly expressive, able to convey aesthetic judgement [3] and other forms of data that resist quantification. Humans are also good at distinguishing between collections of points representing a face, and statistically similar collections that don’t [4]. Since our computers aren’t, they need either very well-fitted manifolds in order to navigate ‘face space’, or a modestly good fit and a human pilot. And if it’s fun for the human to lend their knowledge of face space and its strange boundaries to the computer, then perhaps it’s more fun for them to refine that knowledge co-creatively.

1.1. Portraiture is virtual

Davidenko [5] characterizes various synthetic manifolds (i.e. the expressive range of a generative model) embedded in a (universal) face space equipped with the perceptual distance between any two human faces(2,6),(994,992), and develops a minimal 20-dimensional ‘face space’ of silhouetted face profiles which are suitable as visual stimuli for cognitive experiments. We extend this approach to describe the elaborate ‘face spaces’ which are made explorable by avatar creators in TS3 and similar games, with an eye toward eventually extending the work to face spaces of an illustrative or photographic nature (i.e. lacking a fundamental wireframe representation), that use either
projection techniques (as in life drawing) or texture synthesis (as in generative adversarial networks).

In order to characterize spaces of 3D faces in the context of player creativity, we in fact encounter an intricate pipeline of data structures and percepts. We interpret an output of ‘portraiture’ as encompassing an entire space of possible images of the character represented by an avatar (see Fig. 1, far right), which document the user’s evaluation of the nonliving yet growing character they have represented in the game or fiction.

Simultaneously, we interpret an input of ‘portraiture’ as a particular set of image references (e.g. the face of a person from school, or from a movie, or from another videogame) which the user aims to reproduce, mimic, or modify using the avatar creator. The same artifact may evoke similar or divergent referents in different users, just as the ‘same referent’ may be represented by different artifacts according to different users. An online gallery may easily contain several different Sims of the same character from popular culture.

Both characters and artifacts exist in a semantically rich space. The avatar creator software incorporates certain claims about this space (e.g. ‘eye distance varies within this range’, ‘the height of the inner corner of each eye varies independently of the curve of the lower eyelid’), using a linear basis for additive recombination of facial features, and a rendering algorithm to permit their observation. The immediate visibility of features permits their recombination to be improvised, just as a well-crafted instrument does [6].

1.1.1. Instruments are actual

We choose to study avatar creator software as an instrument in its own right, which encodes part of the skillful practice required to create human likenesses. Their expressive range is only ever partially summarized by their own interface, especially as the numbers of facial feature sliders proliferate, and unpredictable interaction effects [7] start to predominate the output. Even if parameter space is linear, perceptual space is not.

Tubb and Dixon identify the array of sliders with analytic creative interaction, which tends to break down the combinatorial explosion in a face space with many features [8]. In elaborating on our Sims case study, we will see that reflective interaction (i.e. not winnowing, but expanding representations) enters through alternation between gameplay and avatar creation. Conversely, The Sims 4 moves toward tacit interaction - its slider panel GUI is replaced by ‘direct manipulation’ of the Sim face (via well-curated bone sliders, a mapping we discuss in ‘mesh-based geometry’), with panels replaced by ‘coarse’ and various ‘detail’ layers.

To avoid confusion between toys (in the case of small parameter spaces) and tools (in the case of very large parameter spaces, up to and including sculpting software), we call the avatar creation software an ‘instrument’. Like musical instruments, they can be played more or less skillfully. Like analysis instruments, they provide generalizable insight into the character of universal face space.

1.2. The design space of portraiture is vast

Artists learning to draw in the contemporary tradition must perform mimesis of a reference artifact (be it the pose of a model on a stage, stock photos of actors, or even the drawings of
master artists), and are taught to regard this as a productive process. Not only are divergences inevitable anyway, but the process of reaching a landmark in design space is an embodied learning experience [8]. Creating from imagination, then, is just like walking to this landmark across a rugged fitness landscape, - per the navigation-functionality analysis in $9.3$ of [9], - except the artist must recognize the landmark despite its previously unseen form.

Developing a method of documenting forays into the design space of portraiture allows us to not only compare particular results, but also to consider the expressive range of the tool or composer. Reflective tool design [10] makes an artifact of the part of an artist’s skill (often discussed as their ‘style’) that is encoded into e.g. a basis for a certain face space.

Stylization is the visible evidence of a certain set of constraints on the region of design space that a given set of artworks (typically by a closely-knit group of artists) thereby explore. Deliberate constraints are often self-imposed. In Techne Theory (p.188), Staten describes design constraints as ‘maze walls’ in design space, that (at various levels of tangibility, especially the subconscious) define skillful practice itself [11]; a contoured space deeply contingent on material history.

In the realm of avatar creation, these contours may arise either from the tools being used (e.g. the character creator might not have this celebrity’s exact eyebrows), or from particular ways of using those tools (e.g. not using items that didn’t ship with the game, in order to create ‘no-CC’ content). More fundamentally, the contours of universal face space are concealed where they align with the tool’s face space, and revealed where they don’t.

1.2.1. Search methods

We suggest a specific reading of artistic procedures as evolutionary algorithms (after Dennett, via Staten), where the face space of a given creativity support tool operationalizes a single level of abstraction. In particular, sets of landmarks in design space closely resemble the procedure of quality-diversity search [12], which maintains a set of increasingly ‘good’ points in various regions of design space.

It is vital to observe that interactive explorations of design space can be in constant communication with their perceptual landscape (equipped with its rugged fitness function), whereas automatic procedures are better at characterizing computationally intensive properties of the artifacts being designed. Nonetheless, small breaks in this interaction make a design space more interesting, not less. For instance, when a Sim has a child, even if their parameters are more or less evenly combined with their partner’s, the ruggedness of perceptual space can lead to surprises (e.g. if the child inherits a supernatural feature).

Instruments comprise a domain-specific language equipped with direct manipulation, in software terms. Avatar creators are software instruments, and their avatars are similar to Montfort’s ‘player character ... a constraint and possibility defined by the [software’s] author’ [13]. While general-purpose sculpting software can be adapted for use as avatar creators, they produce an experience much less similar to ‘steering’, and much closer to being the vehicle responsible for a safe traversal of design space.
1.2.2. The design space of face geometry is consistently formalized

In this paper, we take particular interest in the language of a linear basis for a face space, that is, some subset of a ‘universal face space’ containing all likenesses of persons.

**Mesh-based geometry**  Meshes consist of a set of points in 3D space, connected by edges forming triangles which discretize a surface. Surfaces that are roughly egg-shaped, but possessing eye sockets and a jawline, can be said to resemble skulls. With the appropriate proportions and elaborations, eventually a face will emerge. Meanwhile, each vertex position should be chosen thoughtfully, to minimize the number of triangles in the resulting surface (to minimize the difficulty of rendering).

We may define a basis for face space as 1) a ‘face mesh’ of this form (including a fixed list of edges between vertices), and 2) a collection of ‘vertex sliders’, each consisting of a linear deformation to the positions of the vertices. In practice, certain subsets of vertices are assigned with variable weights to certain ‘bones’, allowing ‘bone sliders’ which deform the position, scale, and rotation of bones (and ergo those subsets of vertices) to produce more expressive mappings [14]. Because these linear operations compose algebraically, all sliders are basis elements of the face space.

Practical guides to the creation of new sliders (i.e. basis elements in face space) can be found in Sims modding forums [15, 16], relying on a variety of community tools to interface with the game’s data formats.

**Texture maps and lighting**  Diffuse texture maps model the response of each point on the surface to direct illumination, i.e. what color is reflected. Normal texture maps allow bumps and other surface details to be added to the mesh surface without changing the number of vertices.

---

**Figure 2:** Axes in a non-videogame face space developed for use as controlled stimuli in facial perception studies. Note that 2D contours carry enough information to recognize distinct people, and that only three out of 20 axes in the parameterization are shown. (Reproduced from Davidenko, Fig.13.)
The combination of these comprises one possible ‘complexion’ or ‘skintone’.

More advanced texturing and shading techniques, including subsurface scattering (to model the soft translucence of human skin), are prevalent in heavily produced games that invest their resources into ‘realism’ [17]. The relationship of The Sims as a series and of its various communities to this rhetoric is complex. In this paper, we aim to avoid reliance on the intensive character of ‘realistic’ avatar creation, as we are concerned with a separate form of ‘skill’ encoded into a face space.

Yet the evaluation of a face space will necessarily depend upon its rendering. A relatively heavily-produced game such as The Sims (or Skyrim) affords strong curiosity about face space, especially insofar as what needs to be fixed is made evident. Indeed, dissatisfaction with first-party textures can itself drive custom content creation [18] to critique and address the gap.

2. Case Study: The Sims 3 Create-A-Sim

We proceed to consider avatar creation in a modded instance of TS3, looking for moments of expressive payoff resulting from complex technical procedures (such as the slider tutorials discussed above), which distort and remediate the game’s world-model [19]. By considering various iterations of Create-A-Sim (CAS) as a family of modifiable instruments for portraiture, we demonstrate that parameterizations of design space are always tentative.

2.1. Interactive portraiture

2.1.1. Sliders

Sliders are categorized into five parts of the face, and divided among four or more panels in each part. Somewhere between three and a dozen sliders are available in each panel. By default, all of this is hidden, and avatar editing is performed using a gallery of examples. In The Sims 2, the example facial features can be right-clicked to blend toward them by 10%, but this resets the slider GUI for that part’s axes (and not their values), allowing extreme features to be created without mods.

The NRAAS family of community-maintained mods for TS3 (available at nraas.net) implements most of the features we will proceed to discuss, including a litany of bug-fixes and error traps appearing alongside changes to gameplay, and quality of life features such as the addition of most cheat console capabilities into the game UI, and the option to force Create-A-Sim (CAS) to display. In particular, this means that Sim parameterizations (including sliders, skintones, and other non-clothing accessories) need not be locked in, so that extended avatar editing remains a part of the gameplay loop.

**Slider range, and caricature**  NRAAS enables CAS to display an arbitrary number of custom sliders, and can also expand the ranges of all sliders (including those of the base game) by an arbitrary factor - because the value of a slider becomes the floating-point weight of a vector, which is not actually bounded by the suggested range. This is useful for producing humanoid characters with exaggerated or animalistic features, in the style of special effects makeups.
The Monster Factory strategem of ‘no middle sliders’ calls for all sliders to be moved, preferably to an extreme position, regardless of the result. This maximizes the salience of the relationship between any two sliders in the face space, and since these relationships produce the rugged landscape of the space, some of its broad contours can be seen.

For instance, if the eyes are way over there, what will happen to the curvature of the nose bridge area if we lower the brow ridge? We can try it and see. Maybe the brow bone is a child of the eye bone, and the entire silhouette of the brow region changes. Or maybe the brow bone is still where the eyebrows were, and the nose bridge ends up longer.

Adapting Calvino’s terms for a combinatorial expressive space [21, 22], these slider choices define the crystal nature of the face space, acting as authorial choices. They are subject to the clinamen, which is any deviation from pattern that destabilizes the percept by a minimal intervention. In order to locate a potential clinamen, the player-author collaborates with the avatar creator, in conversation with their gameplay intent.

The clinamen in face space depends on the highly nonlinear relationships between sliders, and their ensuing effects on face perception. (Almost all of the points in high-dimensional spaces are in ‘corners’, further from the center than extremal points along any one axis are [23].) For instance, Monster Factory begins with a garishly additive, over-detailed bout of character creation, which promptly lands in an odd corner. These avatars are then juxtaposed with the world, which the game engine happily furnishes their interaction with, since nothing unusual has occurred from its point of view. As a result, suspension of disbelief creeps back in, and characters emerge.
Slider defaults, and vanilla pudding  Certain service roles in TS3 generate new sims with a face that isn’t randomized, but whose sliders are simply all set to zero. The NRAAS mods include both batch processing tools (e.g. by drawing from a gallery of premade Sims) and enhanced console commands for solving the problem (e.g. by loading each affected Sim into CAS, or by remixing them from other Sims in the neighborhood).

The activity of eliminating default faces from the neighborhood involves player community awareness of the face space, which is shared among players through this troubleshooting. It also demonstrates a mode of engagement where NPC avatars are actively expected to be the subject of exploratory interaction (co-creation) with the face space [24].

2.1.2. Interaction with texture maps and accessory slots
While a Sim’s facial geometry is determined by their sliders (relative to the base head mesh), the rest of their appearance depends on skintone, hairstyle, and other bespoke assets.

Combining makeups  Unlike other sculpting software (e.g. Sculptris, 3D-Sculpt) that enables direct painting, CAS provides a variety of makeups (split amongst five slots), with up to four customizable color channels each. This is useful not only for eyeshadow and face-paint, but also for eyebags, ‘contact lenses’ (including both costume lenses, and realistic options using more than one color channel), and any other feature that can be placed into a facial texture.

NRAAS can prevent makeup from taking up its own slot, allowing arbitrary combinations thereof. Clearly, a much smaller number of ‘good’ design options are opened up than bad or pointless ones by this choice. Yet this is a trade-off that certain users are happy to make, which is not surprising - to bother with mods at all means making instrument more tool-like (improving user control), at the cost of making it less toy-like (intensifying user effort).

Iris scale and hat bones  Sliders can also interact with makeup and accessories. For instance, the iris and pupil sliders of #aWT change the radius of these features (regardless of which eye texture the user has applied) by moving the vertices of the eye mesh. Modders have also created sliders for the bones to which things like hats, glasses, and teeth attach, that modify their position to prevent (or accomplish) clipping through the face of a Sim with extreme features.

This represents an extension of face space to accommodate specific textures and accessories. So our story about face space cannot be that it is a nice and rendering-agnostic characterization of expressive range. Rather, it is made into an evolving site of identity work, in its capacity as a small lab of non-narrative world-building activity, as is The Sims as a series [25], as are the communities that have built themselves around it.

3. Conclusions

3.1. Call for face spaces
Linear approximation by sliders is an imperfect but tractable method of modelling universal face space. As an analogy, consider the RGB (red-green-blue) color space, which cuts through the space of all possible spectral signatures (as a submanifold). RGB color space consists of all
possible blends of only three frequencies: a certain red, green, and blue, which each activate their respective receptor in the human eye. But RGB is a poor representation for designers, because the hue and tone of a color varies wildly as red, green, and blue are added or subtracted. A well-chosen linear change of basis will convert RGB space into HSV (hue-saturation-value), or another perceptual color space, which has the same expressive range but is easier to navigate.

Certain ‘explainable AI’ tools are designed to identify human-interpretable bases for expressive linear spaces. For instance, photorealistic 2D face spaces have enjoyed a surge of popularity due to recent work in generative adversarial networks (GANs). The SpaceSheet interface [26] is in fact an avatar creator incorporating real-time user modification of the parameter space, represented as a spreadsheet containing a high-dimensional vector in each cell, which is rendered as a point in the face space of the GAN. Unlike the paper-doll avatar creators commonly available online [27], GANs have not yet seen popular incorporation into feminine play.

Because of the comparative complexity of 3D rendering, libraries specific to avatar creation [28] may be useful in general for creating web applications that rely on morph targets. faceMaker [1] is a well-executed, very plain example of this kind of software. It has a substantial expressive range, but unfortunately, no landmarks besides the default face. Their most characterful slider by far is ‘style’, which produces avatars with the huge eyes and tiny necks of a Disney princess when moved from ‘real’ to ‘cartoon’. It also predominates easily over other sliders. Nonetheless, imagining other ‘style’ sliders (e.g. for assorted interpretations of vampires) is a fun exercise.

3.2. Call for instruments

Instruments are a technology that changes over time. Instruments are defined by their design space, as when sculpting software uses clay as a (visual) metaphor to describe brush-based manipulation of a surface mesh. Yet play with instruments also re-defines their design space, as when situations arising in a life simulation game reflect back on the avatars populating that space, e.g. via the player-author seeking out specific custom content.

Reflective design is a mode of design research oriented toward documenting design knowledge in software instruments and other intelligent interfaces [10]. Avatar creators are instruments of this research when they are oriented toward understanding a face space, meaning that they both make claims (on their author’s behalf) and challenge them (through their users). Two people attempting to use the same instrument to recreate the same person will find surprising differences in those representations.

The reference of avatars to a character who is important to the player, i.e. self-relevance, is a powerful means for players to understand and control their self-perception [29]. Ratan and Dawson provide evidence that when avatar-body disconnection after use is minimized, then avatar self-relevance after use (i.e. during downtime) is maximized. By this argument, indirect control of household Sims (and AI control of self-relevant but non-household Sims) would actually strengthen their connection to the player who isn’t actively playing the game.

However, the strength of self-relevance depends on whether the design space of the expressive instruments in the game, including the face space of the avatar creator, is able to accommodate the player’s creative range. Therefore, even when its face space is not the explicit subject of design inquiry, any life simulation game (e.g. the Sims-like Paralives, or the roguelike Ultima
3.3. Call to create communities

Each instrument affords a common language by which universal face space can be explored, allowing landmarks to be shared in the form of directions to them. Therefore, we assert that the Simming community itself represents the primary good created by a platform which creates both support for novice creators, and room for virtuosity.

Avatar creators are deeply connected to social sharing, arising in the majority of 3D massively multiplayer online roleplaying games [31], as well as in paper-doll avatar creators shared online (e.g. on the website meiker.io). For instance, the Mii Parade that shipped with the Wii console was a space in which the 3D Mii avatars could 'mingle' with Miiis from other consoles [32]. Using the sliders to change the X/Y placement of certain features on the Mii's head, it was possible to create a variety of masked and alien Miiis, e.g. in order to figure out how someone else had entered a certain popular character into a Mii competition.

An intervention to seed this sort of community on other platforms might look like a creative jam oriented around using a particular existing or novel instrument of portraiture. For instance, turnkey stylizers like AI Gahaku [33], which projects selfies into a face space based on Renaissance art, are designed for social sharing. Unfortunately, this kind of face space typically lacks diversity [34], and remediation (e.g. via photo editing) is not afforded in a turnkey design.

As a platform, The Sims facilitates reflective learning [35] thanks in no small part to its maximalism as a life simulator. It supports usage by academics to facilitate skill development and personal expression, using authentic problems such as portraiture [36].

3.4. Acknowledgements

This paper could not have been written without the insight of folks who have partaken in the Casual Creators discussion group, ensuing from the Curating Simulated Storyworlds discussion group. They include Melanie D., Kavi D., Tamara D., Jason G., Isaac K., Max K., Cat M., and many others.

References


