Personalizing Products with Stylized Head Portraits for Self-Expression

Yang Shi
iDVX Lab, Tongji University
Shanghai, China
yangshi.idvx@tongji.edu.cn

Yechun Peng
iDVX Lab, Tongji University
Shanghai, China
2333343@tongji.edu.cn

Shengqi Dang
iDVX Lab, Tongji University
Shanghai, China
dangsq123@tongji.edu.cn

Nanxuan Zhao
Adobe Research
California, USA
nanxuanzhao@gmail.com

Nan Cao
iDVX Lab, Tongji University
Shanghai, China
nan.cao@gmail.com

ABSTRACT

Personalizing products aesthetically or functionally can help users increase personal relevance and support self-expression. However, using non-abstract personal data such as head portraits for product personalization has been understudied. While recent advances in Artificial Intelligence have enabled generating stylized head portraits, these images also raise concerns about lack of control, artificiality, and ethics, which potentially limit their broader use. In this work, we present PicMe, a design support tool that converts user face photos into stylized head portraits as vector graphics that can be used to personalize products. To enable style transfer, PicMe leverages a deep-learning-based algorithm trained on an extended open-source illustration dataset of characters in a cartoonish and minimalistic style. We evaluated PicMe through two experiments and a user study. The results of our evaluation showed that PicMe can help create personalized head portraits that support self-expression.

CCS CONCEPTS

• **Human-centered computing → Interactive systems and tools**; • **Applied computing → Arts and humanities**.

KEYWORDS

Stylized Head Portraits, Artificial Intelligence, Design Support Tools, Self-Expression

ACM Reference Format:
Yang Shi, Yechun Peng, Shengqi Dang, Nanxuan Zhao, and Nan Cao. 2024. Personalizing Products with Stylized Head Portraits for Self-Expression. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3613904.3642391

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642391

1 INTRODUCTION

Product personalization describes a process that changes the appearance or functionality of a product to increase its personal relevance and distinction [35]. The motivations for product personalization are related to the human desire for *self-expression* [29]. For example, many handbag companies allow customers to select the leather type, color, size, and other design components when ordering a bag. The customers can even have their initials or name engraved on a metal badge on the bag. Through this act of self-expression, they can convey their individuality [52, 57], gain positive attention [2, 54], establish social bonds [24], and satisfy creative urges [79].

Researchers in the Human-Computer Interaction (HCI) community have explored various approaches to self-expression by embedding personal data onto the appearance or within the functionality of physical products [3, 11, 22, 23, 35, 62]. For example, Trace-Marker [35] helps engrave patterns on bicycle bags based on a user's history of bicycle-related experiences while LOOP [62], a physical artifact, provides a visualization of a user's step data by changing its behavior. In contrast, exploring self-expression with non-abstract personal data such as head portraits, an important visual representation of individual identity, has received limited attention. Embedding such personal data onto the appearance of physical products raises key considerations around both designing and generating head portraits. First, it is unclear *what* factors should be considered when designing head portraits. For example, the artistic style and emotional tone of a portrait are important - an iconic and fun head portrait will likely appeal to many. Second, *how* to automate the process of designing stylized head portraits remains a challenge. Recent advances in Artificial Intelligence (AI) such as neural style transfer enable rendering specific artistic styles. However, these Generative Adversarial Networks (GANs) [16] or Contrastive Language-Image Pre-training (CLIP)-based models [55] struggle when the domain gap between the source (e.g., illustrations) and target (e.g., face photos) is especially large. Also, current datasets lack sufficient paired examples of illustrations and corresponding face photos to enable models to learn mappings between these disparate domains. Text-to-image models such as Midjourney [47] and DALL·E [51] also support transforming face photos into stylized head portraits. However, lack of control, artificial-looking results, and ethical concerns also spark doubts about the broader applicability of these images.

Motivated by the aforementioned challenges, we present PicMe, a design support tool that facilitates personalizing products with stylized head portraits for self-expression. Specifically, as a user uploads a face photo, PicMe can generate stylized head portraits in vector format based on the input. The system also allows the user to interactively add embellishments to the portraits and preview the effect of the result superimposed on the selected physical product. Our approach to generating stylized head portraits requires an open-source dataset that contains hand-drawn character illustrations in vector format as the training dataset. To this end, we selected the Open Peeps dataset [70] and expanded it by inviting a group of designers to draw more illustrations with diverse facial features and accessories. This extended dataset allows us to design a deep learning (DL)-based algorithm to learn the artistic style of Open Peeps illustrations and transfer it to face photos, thus generating stylized head portraits. We evaluated the effectiveness of PicMe through two experiments and a long-term user study. The results of our evaluation showed that PicMe can effectively generate stylized head portraits that can be used to personalize products and has the potential to improve self-expression.

The main contributions of this work are as follows:

- We present PicMe, a design support tool that helps users personalize products with stylized head portraits for self-expression. Specifically, PicMe allows users to upload a face photo, interactively generate stylized vector portraits with embellishments, and preview the portraits printed on products. The tool can be accessed at https://idvxlab.com/picme/.
- We designed a DL-based algorithm to enable transferring the artistic style of Open Peeps illustrations to source face photos. Our algorithm performs the task in stages to bridge the domain gap between illustrations and photos. We also designed a method for extending the Open Peeps dataset to collect more face illustrations for training our algorithm.
- We conducted an evaluation consisting of two experiments that assessed the performance of our algorithm and a user study that examined the usability of the system. The results showed PicMe was effective at generating high-quality stylized head portraits and encouraging self-expression.

2 RELATED WORK

Our work builds on prior research on self-expression through product personalization and style graphics for head portraits.

2.1 Self-Expression through Product Personalization

Self is described as a "*constellation of multiple, context-dependent identities*" [10]. These identities can be built, maintained, adapted, and reinforced through various means by people to express their "selves" to others [52, 57], which is called self-expression. Product personalization serves as one of the effective means of self-expression [17] and has been applied to various products, including physical products such as sneakers [48] and clothes [36] as well as virtual products such as user interfaces [45, 74] and avatars [39]. In the HCI community, product personalization can be achieved by embedding personal data onto artifacts or installations. For example, Trace-Marker [35] can support engraving patterns on bicycle bags based on a user's bicycle-riding history. Similarly, ListeningCups [11] collects environmental sound data around a user to customize cup patterns. Khot et al. [22] designed SweatAtoms that uses heart rate data to generate 3D printed material artifacts while Nissen [50] encouraged visitors to create personalized souvenirs based on their experiences with cultural events. Also, HCI researchers have explored product personalization by embedding personal data within the functionality of artifacts or installations. For example, LOOP [62], a data sculpture, can rotate to record the number of steps that users take each day. Go and Grow [3] system links individuals' activity data to the amount of water that their personal plant receives. TastyBeats [23], a fountain-based interactive system, utilizes heart rate data from physical activity to create a personalized sports drink. Such design practice involves personal data to encourage users to reflect on and reminisce about their own activities.

In addition to abstract data, non-abstract data such as head portraits can be used to personalize products, as people subconsciously form impressions of personality when presented with portraits [49]. Prior work [73] suggested that greater similarity between users and their avatars can result in more positive attitudes and evaluations towards the avatars. While head portraits are primarily applied to virtual environments, such as video games [27, 75], online communities [13, 46], and virtual reality [15, 59], they are increasingly being used to customize physical products such as photo frames, T-shirts, and jewelry. Our work explores self-expression through product personalization using non-abstract personal data, that is, face photos. To do this, we first explored the attributes for designing head portraits to be marked on physical products and then developed a DL-based design support tool to automate the process of generating stylized head portraits.

2.2 Authoring Stylized Graphics for Head Portraits

Stylized graphics authoring can create expressive visual styles that are artistic and symbolic instead of realistic representations. Various algorithms have been developed to support creating different artistic styles, such as caricature [9, 20], ink painting [80], and pixel art [28, 61], and these algorithms can be classified into graphics- and neural-based approaches [82]. Most of the graphics-based approaches to stylized head portraits simulate the drawing process. For example, Pictoon [7] proposed a two-step stylization process consisting of sketching and stylistic enhancement. The process first estimates the sketch using non-parametric sampling. Then users can interactively change the stroke styles and exaggerate the portrait to produce a personalized cartoon portrait. Le et al. [34] incorporated a set of painting rules, such as the golden ratio rule and the psychological stereotypes concerning different shapes, into their stylization algorithm.

Neural-based approaches leverage neural networks to capture the features of a head portrait and then generate stylized images. Generative models such as GANs are particularly effective in this task. For example, both Pix2Pix [18] and CycleGAN [83] implement style transfer through a generator to output stylized images and a discriminator to optimize the performance of the generator during

Group	Alias	Age	Gender	Occupation	Affiliation
G1	U1	21	Female	Undergraduate student	College of art in a university
	U2	22	Male	Graduate student	College of architecture in a university
	U3	21	Male	Undergraduate student	College of mathematics in a university
G2	U4	21	Female	Undergraduate student	College of design in a university
	U5	21	Female	Undergraduate student	College of design in a university
	U6	23	Female	Graduate student	College of design in a university
G3	U7	21	Male	Undergraduate student	College of design in a university
	U8	32	Female	UI Designer	UI department of a commercial company
	U9	25	Female	Research Assistant	Visualization lab in a university
G4	U10	26	Male	Algorithm engineer	Algorithm department of a commercial company
	U11	21	Male	Undergraduate student	College of mathematics in a university
	U12	28	Female	Journalist	News department of a newspaper office

Table 1: The demographic information of our study participants in the user interview.

the training process. More examples include CariGAN [6], WarpGAN [65], StyleCariGAN [20], and AutoToon [14] have leveraged exaggerated facial deformation and cartoon style transformation for generating head portraits. In addition to generative models, text-to-image models such as CLIP [55] have shown great potential for solving image-related problems due to their capability of evaluating the similarity between images and texts. For example, StyleCLIP [53] allows users to edit head portraits based on textual prompts. The aforementioned models are pre-trained on pixels and thus generate raster images instead of vector graphics, which can limit the application scenarios of the images.

In addition to raster images, CLIP-based methods are able to generate vector graphics. For example, given photos as input, CLIPasso [78] and CLIPascene [77] can generate sketches by using Bézier curves as brush strokes for differentiable rendering [40] and guiding optimization through CLIP. These approaches enable accurate style transfer and improve model robustness through pixel alignment and small distortion. However, they may struggle to effectively transfer the abstract style present in Open Peeps illustrations to face photos due to a domain gap between illustrations and photos. Bridging such a gap may result in a large deformation of facial features when using existing methods. To address the issue, we designed a DL-based algorithm that encodes face photos and illustrations into a shared latent space. This allows PicMe to generate the semantic content of the face photo in the style of the illustration in vector format.

3 PRELIMINARY STUDY

To understand the attributes of head portraits intended for printing on physical products, we conducted formative in-person interviews with 12 users, followed by formative in-person interviews with four professional designers. From both the user interviews and designer interviews, we derived a set of design requirements for PicMe.

3.1 User Interviews

The goal of the user interviews is to gain insights into the following questions: (i) the motivations for printing head portraits on physical products and (ii) the potential challenges presented by such a process of product personalization.

3.1.1 Methodology. We recruited 12 participants (7 females) who ranged in age from 21 to 32 (M = 23.50, SD = 3.43). Their educational backgrounds also vary, including design, literature, mathematics, communication and journalism, and architecture, as shown in Table 1. All participants reported that they have experience in printing their head portraits on personal possessions. We conducted a series of interviews with the 12 participants (**U1-U12**) and the sample questions include "*what type of head portraits did you use to personalize your physical possessions?*" and "*what are the benefits as well as concerns when printing your head portrait on your physical possessions?*" Each interview lasted about 20 minutes.

3.1.2 Analysis and Findings. To analyze the qualitative data collected from the interviews with the participants, we followed a thematic analysis process [4]. Specifically, we transcribed the recordings and then coded the data based on the two proposed questions.

Motivations. More than half of the participants believe that printing head portraits on physical products can **bring enjoyment**, especially cartoon-style head portraits, which are "*cute and fun*" and can make personal products "*more engaging and interesting*". In addition, Four participants mentioned that it can **facilitate self-expression**, as the head portrait can reflect some of her personality and may even influence social interaction. U2 suggested that using head portraits can help **develop a sense of ownership** related to the products. U12 noted that she enjoys having her head portrait serve as a signature added to her possessions.

Concerns. All participants raised concerns when printing head portraits on physical products. First, **information privacy** was mentioned the most frequently by the participants. U6 noted that

Alias	Age	Gender	Occupation	Affiliation	Experience
D1	29	Female	Research assistant	Visualization lab in a university	4 year
D2	51	Male	Professor	College of design in a university	5+ years
D3	26	Female	Product designer	A self-operated design studio	3 years
D4	21	Female	Undergraduate student	College of design in a university	1 years

Table 2: The demographic information of our study participants in the designer interview.

"*I prefer a stylized head portrait that can protect my information privacy but also show my characteristics. In this sense, an artistic head portrait instead of a realistic one is a better choice.*" **Social awkwardness** is another concern related to using stylized head portraits. Many participants felt that head portraits of a realistic style were unattractive to others and showed strong narcissism, using head portraits of cartoonish and minimalist styles instead can avoid embarrassment. The participants also noted that realistic head portraits are less used due to **lack of distinction**. U2 mentioned that using a cartoon-style head portrait can exaggerate certain personality characteristics and reflect personal images in a concise way. E6 argued, "*showing expressions and movements through cartoons is more expressive*".

3.2 Designer Interviews

To understand personalizing products using head portraits from the perspective of professionals, we conducted a series of semi-structured interviews with designers.

3.2.1 Methodology. We recruited four designers (**D1-D4**) by word of mouth, aged between 21 and 51, both with professional experience in design or sketching, as shown in Table 2. In the interview, we asked questions, including (i) what workflow do they adopt for designing head portraits, (ii) what parts of the workflow are easy or difficult to manage with existing tools, and (iii) what guidelines or techniques they use to support printing head portraits on given physical products. Each interview lasted approximately 30 minutes and was recorded for subsequent analysis.

3.2.2 Analysis and Findings. We transcribed the recordings from the interviews and analyzed the qualitative data based on the identified workflow, including designing a head portrait, adding decorative elements to the portrait, and previewing the rendered outcome of the printed portrait on a given product.

Design. According to the designers, creating a head portrait that expressively reflects a user's characteristics is the first step in the workflow. Specifically, individual characteristics are mainly represented in two key aspects. First, designers should effectively capture the user's physical attributes such as facial features and hairstyle. Second, designers should depict the user's unique personality through his or her emotions and create an atmosphere that resonates with the user's character. Insights obtained from the designer interviews also underscore user preference for cartoonish styles over realistic styles. As D4 mentioned, "*users are drawn to head portraits that aren't highly detailed, but have unique, concise, and expressive qualities*".

Embellishing. Once the stylized head portrait has been created, designers are often required to supplement the design with decorative elements. D1 said, "*users usually request additional embellishments for their head portraits, such as festive elements and accessories*". D3 mentioned that "*adding embellishments or gestures would allow the same portrait to communicate different messages when printed within different contexts and on different products.*" In light of their experience, the designers agreed that a design support tool should offer users a diverse choice of embellishments to choose from during the design process. Also, including such embellishments contributes a greater sense of narrative according to the designers. D4 recalled her experience, "*printing one's portrait depicting sports activity onto a tennis racket would evoke an enhanced feeling of immersion.*"

Preview. After decorating the stylized portrait, the designers proceeded to preview it for physical print. As mentioned by D1, users should be able to choose the position and scale of the head portrait printed on a physical product, enabling them to control the final appearance of the product. Similarly, D3 reflected, "*a user's choice of head portraits often depends on the size of the product being customized. For instance, people tend to prefer large, bold patterns on umbrellas, while smaller, more delicate designs are favored for keychains.*" To echo that, all designers expressed a preference for utilizing vector graphics, which are particularly suitable for physical print and fabrication. D2 explained, "*head portraits in vector format can be resized without becoming distorted. This makes them well-suited for printing on products of varying sizes and materials.*"

3.3 Design Requirements

Based on the findings derived from both the user interviews and designer interviews, we established three design requirements to inform the design of PicMe.

- **DR1** Generating stylized head portraits with cartoonish and minimalist styles that effectively capture a user's physical attributes and personality. This will enable users to design expressive portraits that resonate with their personal characteristics.
- **DR2** Providing a diverse library of embellishments that the user can add to his or her head portraits to enhance personal meaning and accommodate to different contexts. Adding relevant embellishments will allow users to create narratives and tailor portraits for different scenarios.
- **DR3** Allowing previewing and adjusting the placement and scale of the head portrait when printing it on physical products. Vector graphics are preferred as they can be resized without

distortion, allowing printing on products of varying sizes and materials.

Our approach to generating stylized head portraits that satisfy the design requirements (**DR1**-**DR3**) requires a vector-based illustration dataset that features a cartoonish and minimalistic style as the training dataset. Thus, we compared different open-source datasets that fall into this category and selected Open Peeps [70] as the base of our training dataset. Open Peeps is an illustration library that includes over 300 hand-drawn characters in various expressions and characterizes a black-and-white hand-drawn style. Note that while only Open Peeps was employed as an exemplar in this work, the proposed method is not limited to a specific dataset. We also extended the dataset with more illustrations of embellishments and facial features through a design workshop and data collection.

3.4 Design Workshop

The design workshop was conducted to explore user preferences when adding embellishments to head portraits.

3.4.1 Methodology. We invited the four groups of participants (G1-G4) from the user interviews to join our online design workshop. We asked each group to use Figma, a web-based collaborative design and prototyping tool, for real-time group brainstorming and design sharing. Also, the participants were required to prepare paper and pens for sketching. Their sketches were photocopied and uploaded to Figma. The workshop began with a 15-minute introduction explaining the goal of our study and the key concepts such as product personalization and style transfer for head portraits. Then, each group was asked to collaborate on three tasks, including i) list physical products that they would like their head portrait printed on (please list each product on a sticky note), ii) categorize the above products based on their usage scenarios. More relevant products can be added to each category (please group the corresponding sticky notes into categories. One product can be in multiple categories), and iii) for each category, draw potential scenes which can be used to embellish the portrait. Scenes are specific moments within a scenario. The workshop lasted approximately 30-50 minutes for each group.

Scenario	Products
Work	Cup, Laptop, Tote bag, Notebook, Folder, Lunch box, USB drive, Schoolbag, Mobile phone case, Sticker, Bottle, cardholder, Pencil case,Key ring, Earphone case, Calendar Cap, T-shirt, Hoodie
Study	Notebook, Tote bag, Folder, Schoolbag, Bottle, Laptop, Pencil case, Name tag, Cup, USB drive, Calendar, Cardholder, Mobile phone case, Sticker, T-shirt, Hoodie, Key ring, Lunch box, Earphone case, Badge, Postcard
Sport	Bottle, Equipment bag, Cap, Tote bag, Hoodie, T-shirt, Sticker, Skateboard, Socks, Umbrella, Badge, Mobile phone case
Home	Cup, Organizer, Fridge magnet, Key ring, Lunch box, Coaster, Bottle, Pillow, Plate, Calendar, Mobile phone case, Earphone case, Cap, Hoodie, Socks, T-shirt
Gift	Cup, Key ring, Bottle, Postcard, Fridge magnet, Sticker, Hoodie, T-shirt, Earphone case, Mobile phone case, Cap, Cardholder
Outdoor	Suitcase, Bottle, Lunch box, Umbrella, Key ring, Name tag, Tote bag, Schoolbag, Postcard, Cap, Hoodie, Socks, T-shirt, Cardholder, Earphone case, Badge, Mobile phone case

Table 3: The six scenarios and corresponding products identified from the design workshop. The products within each scenario are sorted by the number of times they were mentioned by the participants in the design workshop.

3.4.2 Analysis and Findings. By analyzing the four collaborative whiteboards collected from the design workshop (Figure 1), we found that the decorative scenes drawn by the participants have the following characteristics. First, **six distinctive scenarios of scenes were identified, including work, study, sport, home, gift, and ourdoor**, each with a set of common products, as shown in Table 3. Second, **the scenes can be categorized into activity scenes and state scenes.** Specifically, activity scenes show that the person is actively engaged in a specific task, such as reading and giving a presentation while state scenes convey the lifestyle and inner thoughts of the person, such as feeling energetic and slacking off. Analysis of the participants' whiteboards also revealed a higher proportion of state scenes in the study, work, and gift scenarios, while activity scenes were more prevalent in the sports, home, and outdoor scenarios. Third, **the participants exhibited a preference for state scenes.** Such a preference can be attributed to the potential of state scenes to convey the person's mood or feelings, which effectively stimulates empathy. Also, state scenes often include more personalized content and interesting ideas when compared with activity scenes.

3.5 Data Collection

Based on the findings derived from the design workshop, we then extended the dataset by recruiting professional designers to draw illustrations of decorative scenes and facial features that can be added to the Open Peeps dataset.

3.5.1 Designer Recruitment. We recruited three out of four designers from the designer interviews. We checked their sketching portfolio to ensure that their proficiency in drawing illustrations and their artistic styles were similar to that of Open Peeps. Each designer worked for 50 hours over two weeks and was compensated 10 USD per hour.

3.5.2 Drawing Illustrations. The designers were required to draw decorative scenes based on the six identified scenarios. Specifically, we first identified the five most frequently sketched and well-received scenes within each scenario. Then, we asked each designer to choose two scenarios with ten scenes and draw illustrations according to the keywords provided and the scenes specified by the participants in the design workshop. The requirement is that the illustrations should precisely express the meaning of the keywords

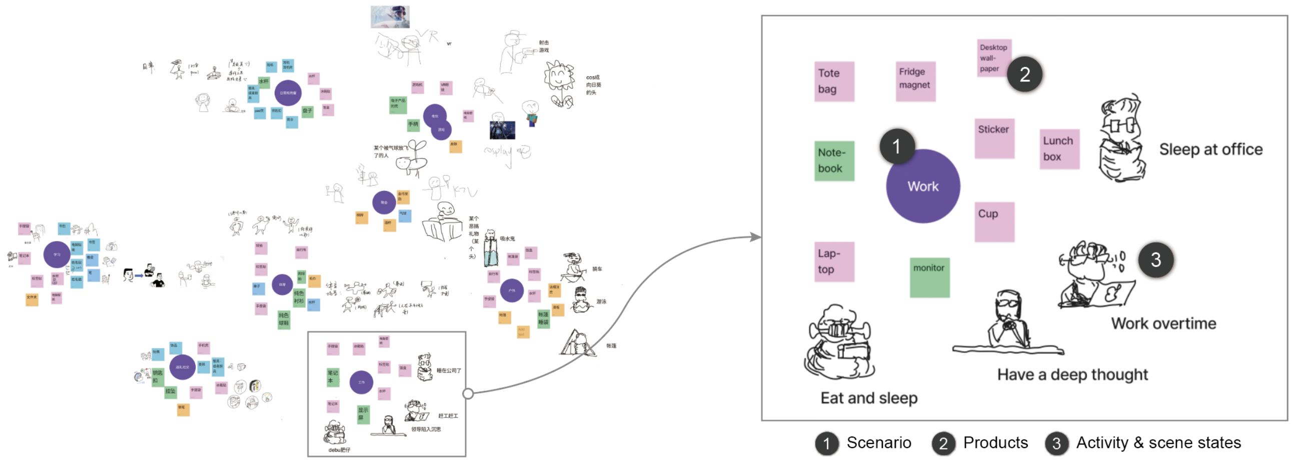


Figure 1: The collaborative whiteboard by Group 3 (G3). In this whiteboard, G3 grouped the sticky notes representing products into eight clusters representing usage scenarios.

without ambiguity, as well as avoid cultural differences. Finally, the designers collaborated with two researchers to review and refine the content and style of the illustrations. Figure 2 shows the final illustrations of 30 scenes drawn by the three designers. Note that the portraits of characters in each scene can be replaced with portraits of other characters in Open Peeps.

Figure 2: The illustrations of 30 scenes categorized by the six scenarios, each with corresponding keywords.

Also, to ensure the diversity of our dataset, the designers were asked to extend it by expanding upon the existing component categories already present within the data, including hairstyles, facial expressions, glasses, and beards (Figure 3). For example, to cover as many common hairstyles as possible, we referred to the hairstyles observed in CelebFaces Attributes Dataset (CelebA) [43] and summarized four characteristics of hairstyle, including length, curliness, bangs, and tying. Note that hair color is not considered as the illustrations of Open Peeps are mostly in black and white. We also delineated characteristics related to facial expressions, including facial features (i.e., eyebrows, eyes, nose, mouth) and emotions (i.e., neutral, happy, sad, angry, fear, surprise, disgust). Next, the designers were required to draw illustrations by referring to these characteristics. Finally, one researcher checked the result to ensure that the extended dataset could cover a random sample (1%) of the face photos in CelebA.

In total, we collected 28 hairstyles, eight facial expressions, eight glasses, eight beards, and 30 scenes from the three designers. These

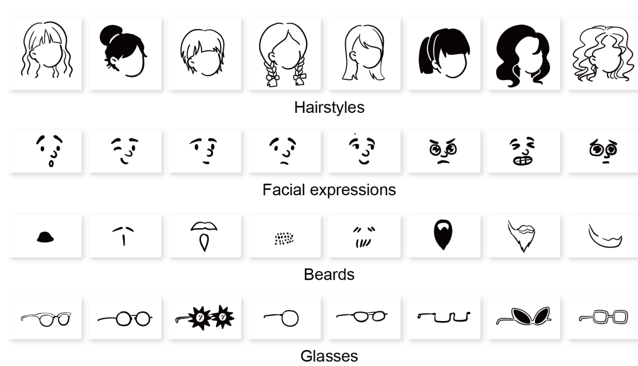


Figure 3: Example illustrations of hairstyles, facial expressions, beards, and glasses drawn by the designers.

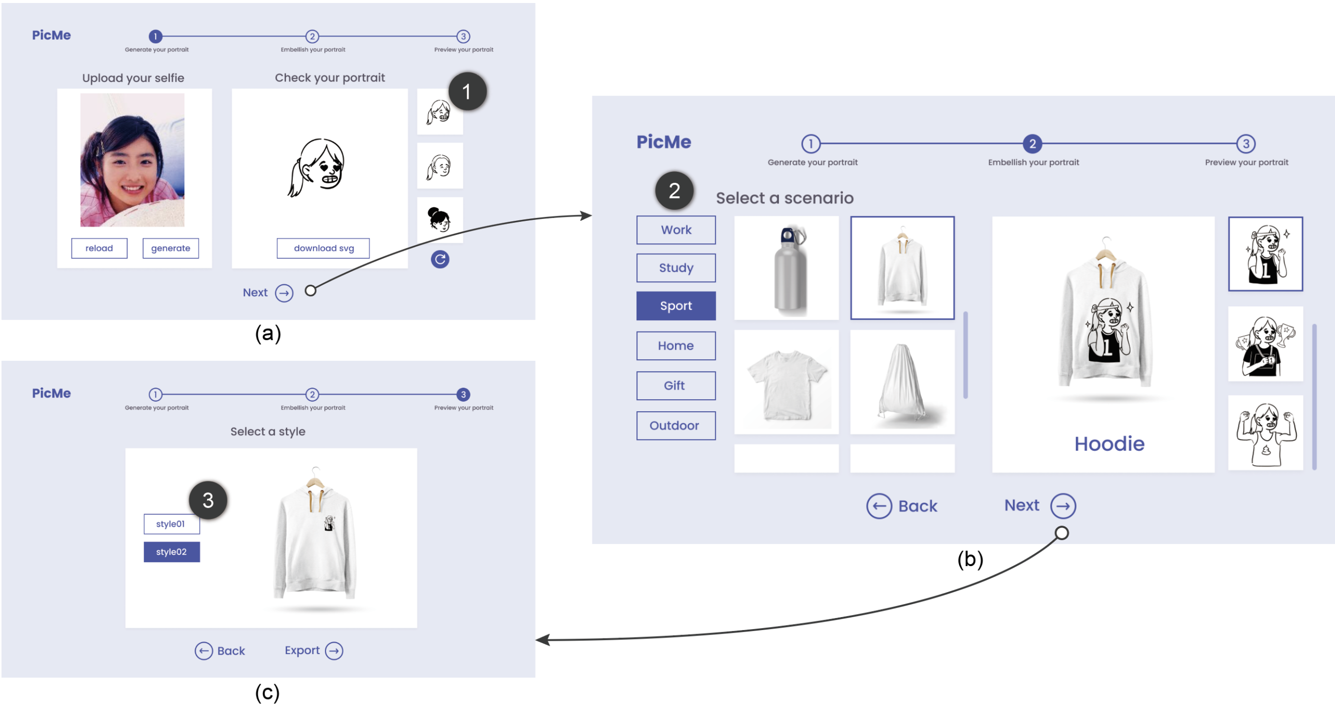


Figure 4: The user interface of PicMe contains three pages, including (a) the Portrait Generation Page, (b) the Scene Selection Page, and (c) the Product Preview Page.

illustrations were sketched with a stylus and tablet in Affinity Designer.

4 PICME SYSTEM DESCRIPTION

In this section, we introduce PicMe, a design support tool that helps generate stylized head portraits that can be used to personalize products and facilitate self-expression. Specifically, we describe how each part of PicMe that meets the design requirements (**DR1-DR3**) described in Section 3, through the details of the user interface, interactions, and algorithms.

4.1 User Interface

The user interface of PicMe is composed of three pages, the Portrait Generation Page (Figure 4 (a)), the Scene Selection Page (Figure 4 (b)), and the Product Preview Page (Figure 4 (c)). The Portrait Generation Page provides the Suggestion Panel (Figure 4 (a): 1) to display stylized head portraits suggested by PicMe. In the Scene Selection Page, the Options Panel (Figure 4 (b): 2) lists options for specifying a product for printing the resulting head portrait as well as a scene for embellishing the portrait. The Product Preview Page contains the Style Panel (Figure 4 (c): 3), which allows users to change the print rendition such as the size, rotation, and location of the portrait.

4.2 System in Action

We describe a representative usage scenario focused on PicMe's core capabilities. This scenario walks through how a user creates a stylized head portrait and transfers it to a physical product, highlighting key interactions and features provided by PicMe.

Generation (DR1). Imagine that a user named Alice attempts to print her head portrait, which serves as imagery that represents herself, on her sports possessions to show her characteristics. Initially, Alice uses her camera to capture a face photo, which is then uploaded to the Portrait Generation Page (Figure 4 (a)). If she is dissatisfied with the photo, she can use a reset function to re-upload an alternative. Upon uploading the photo, Alice can click the 'generate' button on the Portrait Generation Page to instigate the system to synthesize a series of head portraits based on the source. Alice can select one satisfactory result to constitute her stylized head portrait or click the 'update' button to invoke additional iterations to regenerate more results. Upon completion, Alice can click the 'next' button to go to a subsequent page.

Embellishing (DR2). On the Scene Selection page (Figure 4 (b)), Alice chooses an appropriate scenario to complement her stylized portrait. First, she envisions potential real-world applications for her portrait, sport, and clicks the corresponding button in the Options Panel. Next, Alice chooses her preferred product and scene displayed on the panel under sport. She selects sport-affiliated apparel from the given products and determines that the 'full of energy' scene can better reflect her psychological state during sports activities. Upon finalizing the selection of the product and scene, she clicks the 'next' button to go to the subsequent page.

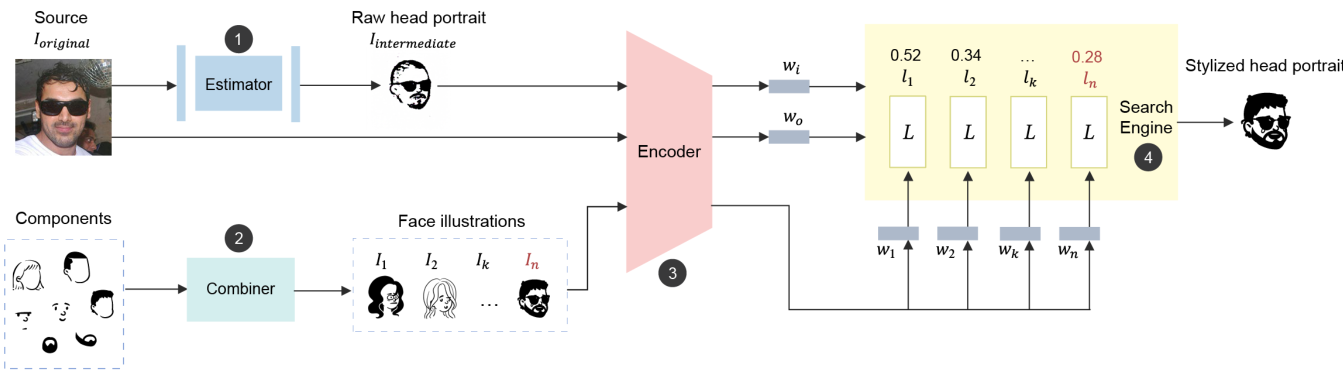


Figure 5: The pipeline of our algorithm, including (1) an estimator, (2) a combiner, (3) an encoder, and (4) a searcher engine.

Preview (DR3). On the Product Preview Page (Figure 4 (c)), PicMe renders a preview depicting the contextually framed portrait situated upon the selected product across various scales and positions. Upon determining the final preview configuration, Alice clicks the 'export' button, which outputs two image files – a concept product image in PNG format along with a stylized portrait in SVG format. Referencing the product image, Alice then inputs the portrait into a surface printing device to print it on her hoodie, completing the process of product personalization.

4.3 Search and Generation Algorithm

Given a face photo uploaded by the user, the objective of the algorithm is to combine the content of the source with the artistic style of Open Peeps illustrations. However, performing such style transfer poses two challenges. First, there exists a significant domain gap between face photos and face illustrations, as illustrations are more abstract than photos. Thus, correctly aligning attributes such as facial features (e.g., eyes) and accessories (e.g., glasses) between the two distinctive domains is challenging. Second, compared to large-scale face photos, face illustrations are far fewer in quantity. Thus, it is difficult to construct a large-scale paired photo-illustration dataset as the training dataset for AI models. To address these challenges, our algorithm integrates an intermediate stage by first converting the photos into illustrations to roughly estimate their features in the domain of illustrations. These features are utilized to bridge the domain gap. Also, we leveraged a pre-trained CLIP model to encode both photos and illustrations into the same latent space, addressing the issue of limited training data.

Thus, we designed an algorithm that is composed of four modules (Figure 5): (1) an estimator that converts the source face photo into a raw head portrait bitmap, (2) a combiner that combines the components from the Open Peeps dataset as face illustrations and renders them as face illustration bitmaps, (3) an encoder that maps the source face photo and all the bitmaps into the latent space, and (4) a searcher engine that finds the most matching face illustrations to the source face photo in the latent space and recommends them to users.

Estimator. The estimator is used to provide a rough estimate of the cartoonish features included in the source face photo by generating a raw portrait bitmap. In this process, we followed the pix2pix [19] method and implemented the estimator by training a neural network, where image translation was achieved using the U-net [58]. We trained the estimator on 720 manually labeled photo-illustration pairs. Specifically, we first randomly selected 720 face photos from the CelebA dataset [43]. Then, for each of the face photos, one out of three designers who were recruited to extend the dataset manually selected components from our dataset that closely match the characteristics of the face photo and combined them as a face illustration. Here, the estimator is denoted as G and represented as follows,

$$I_{intermediate} = G(I_{original}, z) \tag{1}$$

where $I_{original}$ denotes the source face photo, z is a random noise vector, and $I_{intermediate}$ denotes the raw portrait bitmap generated by estimator G.

Combiner. The combiner aims to combine the components from the dataset as face illustrations and render the results as bitmaps, which will be further used in the encoding process. In our dataset, the categories and the number of components in each category are as follows: hairstyles (70), facial expressions (36), glasses (17), and beards (25). We first standardized components by category through affine transformations, including translation and zooming. After standardization, we selected one component from each category and combined them in a predefined sequence, which resulted in more than one million face illustrations. The composite face illustrations are then rendered as bitmaps, denoted as $\{I_1, I_2, ..., I_k, ..., I_n\}$. Here, n is the total number of component permutations ($n = 1,071,000$) while k is the index of every unique permutation. I_k can be represented as a sequence of components, a bitmap, or a vector image.

Encoder. The encoder encodes the source face photo $I_{original}$, the raw portrait bitmap $I_{intermediate}$, and face illustration bitmaps $\{I_k\}$ into the same latent space for subsequent search. Specifically, we utilized the pretrained image encoder of Contrastive Language-Image Pre-Training (CLIP) [56] to encode the images to latent codes and used the pretrained image embedding space of CLIP as the latent space. CLIP was selected as it is a neural network trained on large-scale text-image pairs to align text and images into the same embedding space through a text encoder and an image encoder. By computing the similarity between the latent codes in the image embedding space, CLIP is able to find similar features between the source face photo and the face illustrations.

Search Engine. The search engine is used to find a set of face illustrations from $\{I_1, I_2, ..., I_k, ..., I_n\}$ that best match the source face photo in the latent space. During the process, the face illustration with the minimum loss is selected and recommended to users. When designing the loss function, we considered using the raw portrait bitmap $I_{intermediate}$ with salient facial features to bridge the domain gap. Thus, the loss function is defined as follows:

$$L(w_o, w_i, w_k) = C - cos(w_o, w_i)^p cos(w_k, w_i) - (1 - cos(w_o, w_i)^p) cos(w_k, w_o) \quad (2)$$

while the search goal is:

$$\arg\min_k L(w_o, w_i, w_k) \quad (3)$$

where w_o denotes the latent code of $I_{original}$, w_i denotes the latent code of $I_{intermediate}$, and $\{w_1, w_2, ..., w_n\}$ denotes the latent code of $\{I_1, I_2, ..., I_n\}$. Here, w_o, w_i and w_k are all 512-dimensional vectors encoded by CLIP. The terms $cos(w_k, w_i)$ and $cos(w_k, w_o)$ respectively evaluate the cosine similarity between $I_{intermediate}$ and I_k, and between $I_{original}$ and I_k, balanced by two coefficients, $cos(w_o, w_i)^p$ and $(1 - cos(w_o, w_i)^p)$. A larger value of $cos(w_o, w_i)^p$ suggests that $I_{intermediate}$ can replace $I_{original}$ during the search process, resulting in a larger contribution of $cos(w_k, w_i)$ to the loss function. A larger contribution of $cos(w_k, w_o)$ to the loss function occurs when $cos(w_o, w_i)^p$ is smaller. The hyperparameter p is used to balance the contributions of different components of the loss function and we set $p = 2$. C is a constant to ensure the loss is non-negative. To enable an efficient search, the face illustrations $\{I_k\}$ are pre-encoded and their latent vectors were combined into a $n \times 512$ matrix M, where each row represents the latent code for a specific permutation indexed by k and each column represents a specific dimension of the latent code. Also, the tensor calculation was used to enable parallel computation of the loss.

5 EXPERIMENTS

We performed two experiments to validate the effectiveness and generalizability of our algorithm, respectively.

5.1 Experiment I: Effectiveness

In terms of effectiveness, we evaluated the resemblance between a given face photo and the resulting stylized head portrait through an ablation study. The ablation study replaced three out of four modules of the algorithm, including the estimator, the encoder, and the search engine, to understand the contribution of the modules to the overall performance of our algorithm. Each of the three modules was replaced with one or more baselines, respectively. For all baselines and PicMe except the baselines of the estimator, we set the batch size to 16 and the number of epochs to 200. They were trained using Adam [26] optimizer, with $lr = 0.0002$, $\beta_1 = 0.5$, and $\beta_2 = 0.999$. For the baselines of the estimator, we followed the experimental settings described in [9, 25]. All baselines and PicMe were trained on the 720 labeled photo-illustration pairs from CelebA and Open Peeps, respectively. The ablation study was conducted on a laptop with Intel (R) Xeon (R) Gold 6148 CPU, TESLA V100 PCIe-16GB GPU.

5.1.1 Baselines. Regarding the baselines, we First, we replaced pix2pix in PicMe's estimator with JoJoGAN [9] (denoted as baseline $EM_{JoJoGAN}$) and U-GAT-IT [25] (denoted as baseline $EM_{U-GAT-IT}$), respectively. JoJoGAN employs GAN inversion to acquire training data and fine-tunes StyleGAN to perform style transfer on face photos based on a given style image. U-GAT-IT utilizes the attention module and AdaLIN function to construct a deep learning model for image-to-image translation. For both baselines $EM_{JoJoGAN}$ and $EM_{U-GAT-IT}$, other modules of the algorithm except the estimator remain the same to PicMe. Second, we replaced CLIP in PicMe's encoder with Bootstrapping Language-Image Pre-training (BLIP) [38] (denoted as baseline EC_{BLIP}) to change the image embedding space. Similar to CLIP, BLIP is a language pre-training framework, but it incorporates a Captioner which generates image captions and a Filter which removes noisy captions to utilize web data for training. Third, we used different loss functions in PicMe's search engine, that is, only the information of $I_{original}$ (denoted as baseline $SE_{original}$) or $I_{intermediate}$ (denoted as baseline $SE_{intermediate}$) was applied. Also, we tested the loss under different hyperparameters (denoted as $SE_{p=1.0}$, $SE_{p=1.5}$, $SE_{p=2.0}$, and $SE_{p=2.5}$).

5.1.2 Validation Dataset. We constructed a validation dataset for our experiment. We carefully selected 20 face photos from the CelebA-HQ dataset [21] which were not used in the training process. The 20 face photos cover diverse attributes of people such as age, gender, and skin tone, as well as characteristic features such as glasses, beards, and facial expressions. We conducted a survey to identify the components in the dataset that can accurately capture the features of the face photos. Specifically, we recruited a group of 20 professional designers (4 females) as the participants. They ranged in age from 18 to 27 (M = 21.20, SD = 1.99). The survey required the participants to check a given face photo and rate all the components by category on their resemblance to the corresponding feature present in the face photo using a 5-point Likert scale. Each participant was presented with five face photos, one at a time. As a result, for each face photo, we obtained an average resemblance score for each component to it. The face illustration composed of the highest average scoring components from each category is regarded as the head portrait most similar to this photo.

5.1.3 Measurement. We considered the process of searching the most similar combination of components for a given face photo as a retrieval task, thus we adopt the Mean Average Precision (MAP) [76] as the measurement, which is commonly used in retrieval tasks [12, 41, 81]. MAP is computed as follows:

$$\frac{1}{J} \sum_{j=1}^{J} \frac{1}{\sum_{i=1}^{T} r_{ij}} \sum_{i=1}^{T} (r_{ij} \frac{\sum_{k=1}^{i} r_{kj}}{i}). \quad (4)$$

where J is the total number of tasks and T is the maximum number of returned results. In our task, J can be regarded as the total number of face photos in the validation dataset, that is, 20, while T refers to the maximum number of retrieved face illustrations from the dataset. r_{ij} denotes whether the i-th result returned by the j-th retrieval task is relevant. A higher MAP score indicates a better result. Here, r_{ij} is computed as follows:

$$r_{ij} = \begin{cases} 1, & if \quad s_{ij} \geq c \\ 0, & if \quad s_{ij} < c. \end{cases} \quad (5)$$

Figure 6: The stylized head portraits generated by our algorithm and the baselines. Specifically, baselines $EM_{JoJoGAN}$ and $EM_{U-GAT-IT}$ replaced PicMe's estimator with JoJoGAN [9] and U-GAT-IT [25], respectively. Baseline EC_{BLIP} replaced PicMe's encoder with BLIP [38], baselines $SE_{p=1.0}$, $SE_{p=1.5}$, and $SE_{p=2.5}$ tested PicMe's loss under different hyperparameters. Baselines $SE_{original}$ and $SE_{intermediate}$ were different loss functions in PicMe's search engine. NN stands for nearest neighbor.

where s_{ij} is the resemblance score between the i-th returned face illustration and the j-th source face photo. c is a threshold constant.

5.1.4 Results and Analysis. Figure 6 shows the results generated by the baselines and our algorithm. We found that our algorithm can generate results with similar features to the sources. In comparison, the baselines changing the estimator ($EM_{JoJoGAN}$ and $EM_{U-GAT-IT}$) or using only $I_{original}$ in the search engine ($SE_{original}$) can significantly affect the features of the generated portraits, returning results that are mostly of poor quality. The baselines changing the encoder (EC_{BLIP}), adjusting the value of p in the search engine ($SE_{p=1.0}$, $SE_{p=1.5}$, and $SE_{p=2.5}$), or using only $I_{intermediate}$ in the search engine ($SE_{intermediate}$) can affect the results in a subtle way, returning some results that are of acceptable quality. Figure 7 shows the MAPs of PicMe and the baselines at c = 3, 3.3, and 3.5 for $1 \leq T \leq 10$ as line charts. Also, we added the performance of a method that randomly selects components from the dataset to constitute head portraits to each chart to facilitate comparison. We observed that among different estimators, pix2pix (PicMe) emerges as the optimal choice, especially at higher thresholds c. Across different encoders, CLIP (PicMe) and BLIP demonstrate varying performances at different thresholds, with CLIP's latent space outperforming BLIP's at higher thresholds c. When considering different loss functions in the search engine, the loss generates the best outcomes at $p = 2.0$ (PicMe).

5.2 Experiment II: Generalizability

The generalizability of the algorithm was evaluated by its performance on different datasets. We used the same validation dataset as in Experiment I.

5.2.1 Datasets. We chose Avataaars [71] and Avatar Illustration System [33] as alternative datasets, both of which are open-source and of a minimalist and cartoonish style. When compared to Open Peeps, the two datasets are in color instead of black-and-white, thus encompassing more details. To combine components from these datasets as face illustrations, we utilized avataaars-generator [42] and vue-color-avatar [37] as combiners, respectively. We selected 720 face illustrations from each of the two datasets as the training set, the same as Open Peeps. We trained our algorithm independently on these datasets using the same experimental settings as for the OpenPeeps dataset.

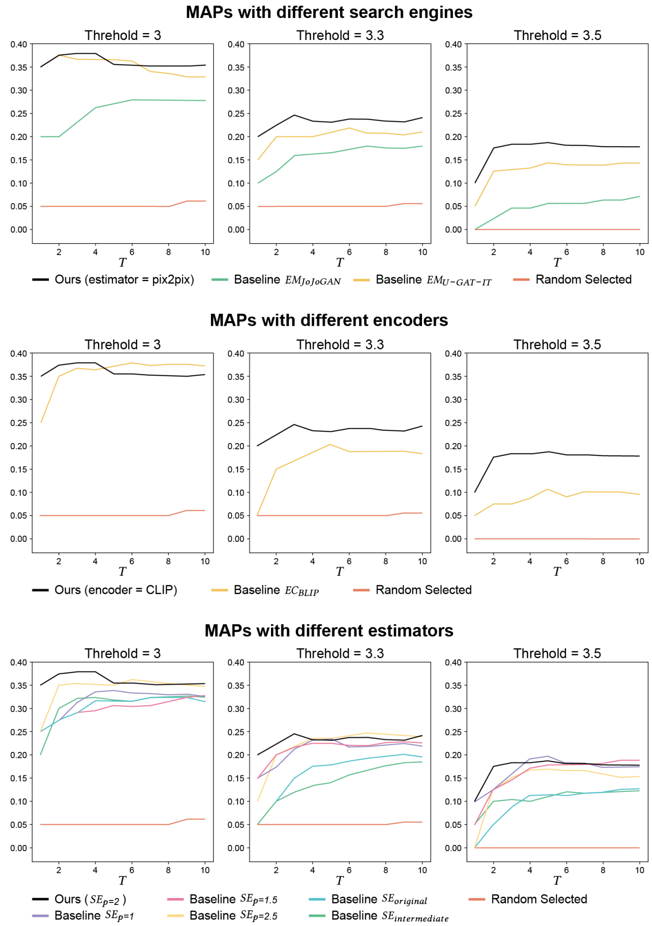


Figure 7: The MAPs of our algorithm and the baselines. T **denotes the maximum number of retrieved face illustrations from the dataset.**

5.2.2 Results and Analysis. Figure 8 shows the results generated by our algorithm using the two alternative datasets. We observed that for face photos possessing varied attributes, our algorithm effectively matches components, including hairstyles and facial expressions. More importantly, our algorithm also pays attention to the color-related features of source face photos without specialized adaptations for skin tone or hair color and generates reasonable results. For example, when comparing different lines in Figure 8, we found that although skin tone or hair color depth varies, the outcomes can demonstrate relatively robust performance.

6 USER STUDY

To evaluate the effect of PicMe on supporting product personalization and augmenting self-expression, we conducted a user study.

6.1 Participants

The participants were recruited through a public exhibition. The research team set up a booth at a community exhibition for 15 days to showcase the work, as shown in Figure 9. Informational posters introduced the work while on-site computers provided interactive demonstrations, inviting the use of PicMe. Over the course of the exhibition, 40 participants (13 males) were recruited. The participants ranged in age from 15 to 30 (M = 22.38, SD = 3.48), representing diverse fields of expertise, such as mathematics, literature, music, computer science, and design. Their experience in design (novice: 45.0%, advanced beginner: 20.0%, competent: 20.0%, proficient: 12.5%, professional: 2.5%) and in product personalization (novice: 60.0%, advanced beginner: 17.5%, competent: 15.0%, proficient: 7.5%) also vary.

6.2 Task and Procedure

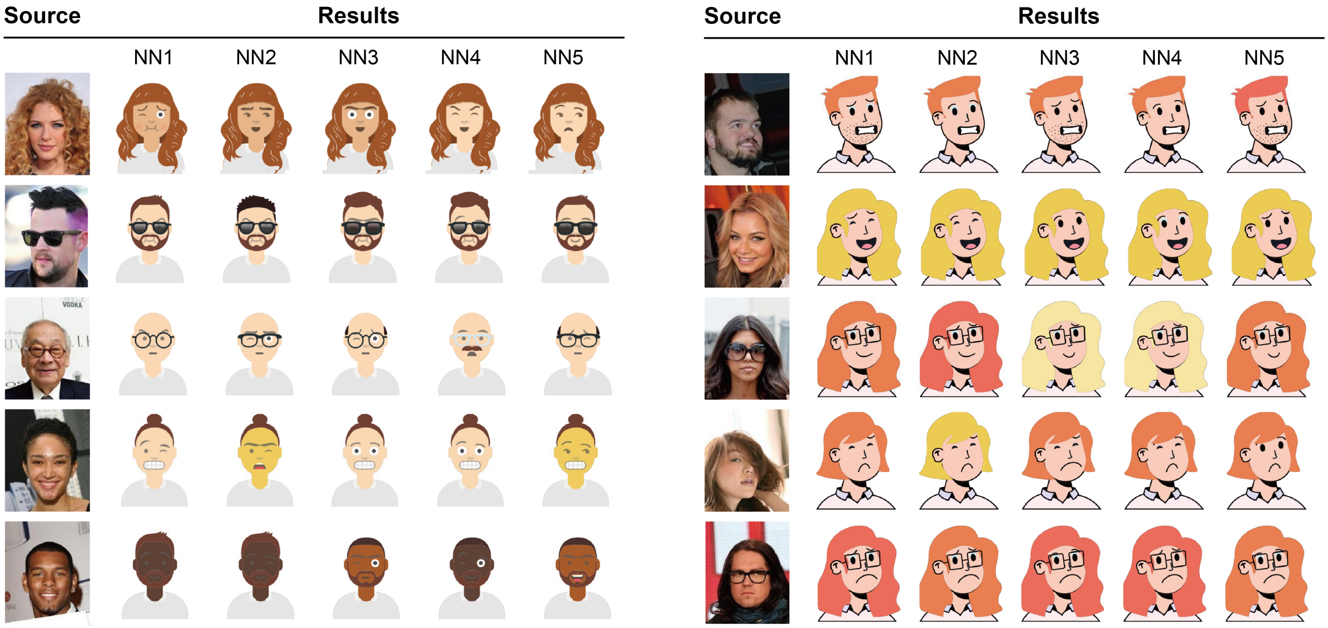
The user study consists of three sessions: the *design*, *fabrication*, and *usage* sessions. In the design session, we started with a 10-minute introduction explaining the goal of our study and obtaining their consent to video recording. Then, the participants were presented with PicMe and instructed to generate stylized head portraits using their own face photos. After the participants finished exploring PicMe, they were asked to rate it using a questionnaire and provide their demographic information. Then, we conducted a semi-structured interview with the participants to learn about their user experiences and the reasons for their ratings. The design session lasted about 35 minutes.

After the design session, we collected 60 product-portrait image pairs designed by the 40 participants and proceeded to the fabrication session. Within five working days, we completed the task of printing portrait graphics on their associated products, totaling 60 personalized products. To accommodate individual needs and constraints, the participants were allowed to mark their stylized portraits onto their own used belongings or utilize fresh merchandise supplied by the researchers. In this session, we utilized a range of printing techniques to print portraits on products with diverse materials. For paper-based items such as notebooks, postcards, and stickers, we used inkjet printers that can precisely print onto flat surfaces. For fabrics such as clothes, schoolbags, and tote bags, we had fabric printers that embedded inks directly into the textiles. For hard surfaces including coasters, cups, and badges, we employed surface printers that can coat images onto ceramic, glass, and metal items. Finally, for materials such as leather wallets and wooden keychains, we had laser engravers that can etch portrait graphics into the materials through laser incisions. These personalized products were then postal mailed to the respective participants as compensation for their contributions (Figure 10).

Six months after the participants received the products, we sent emails to them to investigate their usage experiences and ask about their willingness to take part in a follow-up interview. As a result, all the participants responded positively and we asked them to fill out another questionnaire regarding their use of the personalized products. Then, we randomly invited 20 of 40 participants to join the usage session. In this session, we conducted a semi-structured interview. The session lasted about 0.5-1 hour for each participant.

6.3 Measurements

In the design session, the questionnaire includes questions about (i) the quality of resulting head portraits and (ii) the effectiveness of our PicMe system. For (i), we assessed the *similarity* between



(a) (b)

Figure 8: The head portraits generated by our algorithm using (a) the Avataaars dataset [71] and (b) the Avatar Illustration System dataset [33], based on the source photos from CelebA-HQ [21]. NN stands for nearest neighbor.

the resulting head portraits and the source face photo, the *representativeness* of the portrait (i.e., it reflects or fits one's identity), and its *expressiveness*. For (ii), we measured the usability (*usefulness, efficiency, ease of use, ease of learning*, and *satisfaction*) [1, 44, 63] and creativity support (*results worth effort, enjoyment, exploration*, and *expressivity*) [8, 64] of PicMe.

In the usage session, the questionnaire includes questions regarding (i) user attitude toward product personalization and (ii) user experience of self-expression when using personalized products. For (i), we asked if the participants think product personalization is meaningful (*meaningfulness*). For (ii), we selected items for our questionnaire from the criteria commonly used for self-expression [60, 72], including reflecting individual identity (*identity*), aligning self-image (*self-image*), gaining attention (*attention*), and strengthening social bond (*social bond*).

(a) (b)

Figure 9: The setting of our user study. (a) Our booth at the exhibition. (b) A participant is using PicMe.

6.4 Analysis and Results

Based on the analysis of user study data, we now report both the quantitative and qualitative results. Figure 11 shows the quantitative results of the questionnaire rated by the participants. For qualitative data analysis, we employed the thematic analysis method [5] to analyze the data. The goal is to understand the participants' comments and highlight their thoughts on the quantitative results.

6.4.1 Quality of Results. Overall, PicMe performed well in the aspect of generating high-quality stylized head portraits. Specifically, it received positive ratings regarding *similarity* ($MD = 3.50, IQR = 1.00$; Q1), *representativeness* ($MD = 4.00$, $IQR = 1.00$, Q2), and *expressiveness* ($MD = 4.00$, $IQR = 1.00$, Q3). The majority of the participants agreed that the resulting head portraits are similar to their uploaded face photos. P9 said: "*I'm always putting my hair up in a ponytail. The ponytail in this head portrait makes it obvious that's supposed to be me. It captured my usual style perfectly.*" P4 commented, "*My girlfriend said this pattern totally reminds her of me. The hairstyle and the beard is very representative.*" Also, more participants described their head portrait as "*expressive and interesting*", as PicMe can generate head portraits that reflect both the participants' visual appearance and their personality traits and emotions. P7 said, "*It matched me with a big hearty laugh as my*

Figure 10: Example personalized products designed by the participants, P1-P16, in the user study. The stylized head portraits printed on each of the products were generated using PicMe.

expression. That's perfect since I'm pretty much always smiling and chuckling in daily life. The cheerful vibe of this portrait aligns well with my lighthearted personality." P12 chose 'slacking off' as the scene to match her face, "*This scene totally captures how I feel like spacing out at work sometimes. It's super relatable and expresses that feeling perfectly.*"

Moreover, the participants had a strong sense of identification with the generated head portraits. As P16 said, the head portrait acts as "*a stamp, a symbol of myself*". The majority of participants felt the head portrait could represent themselves, and personalized products with those head portraits would be one-of-a-kind to them. P19 commented:"*I picked this avatar myself and it looks just like me, so I feel like it totally captures the real me. Putting this avatar on my possessions is like stamping it with my own unique seal - it makes my possessions especially connected to me.*" P5 explained: "*This portrait mirrors me and I see myself when I look at it. If it's on my product, I feel like others would be able to recognize my possessions too.*"

6.4.2 Perceived Effectiveness of PicMe. In terms of the effectiveness of the system, PicMe was rated positively regarding *usefulness* ($MD = 4.00, IQR = 1.00$, Q4), *efficiency* ($MD = 4.00, IQR = 2.00$, Q5), *ease of use* ($MD = 5.00, IQR = 1.00$, Q6), *ease of learning* ($MD = 5.00, IQR = 1.00$, Q7). 60% of the participants rated 5 for ease of use and ease of learning, showing the user-friendliness of the system. The participants appraised PicMe for its capability to support product personalization in a user-friendly manner. As P7 mentioned, the system is "*concise, intuitive, and with decent functions*". P5 said, "*The UI design is pretty cool in my opinion, and everything is nicely streamlined and easy to understand. The workflows are clear-cut rather than convoluted, making this a very user-friendly system.*" P1 said, "*The system is quite simple and easy to operate. I think I'd be able to work through the task smoothly without sitting through any tutorials.*" Also, PicMe can provide users with various options to help them create personalized products. P4 felt that the system allowed her to complete the design task in a "stress-free" manner, "*It gives me dozens of choices to pick from, with all kinds of different looks and feels. I can just browse around and choose whatever fits my personal taste.*" P19 added: "*If I have to design something totally from scratch, I'd be clueless since I'm not experienced. This system does all the heavy lifting by giving me a big selection to choose*

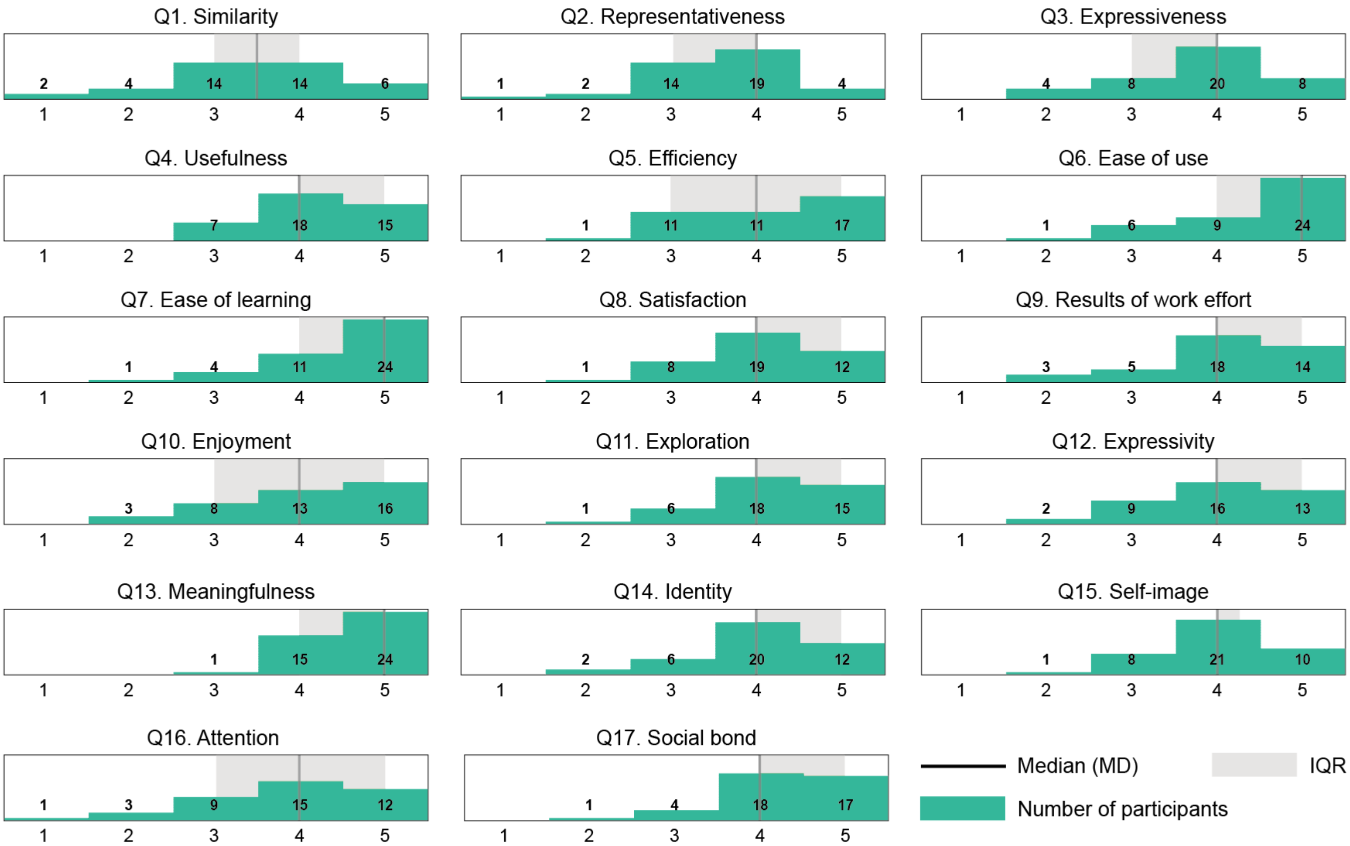


Figure 11: The participants' responses to the quality of resulting head portraits (Q1-Q3), the usability and creativity support of PicMe(Q4-Q12), product personalization (Q13), and self-expression (Q14-Q17). Each participant answered on a 5-point likert scale with the questionnaire.

from, so I can quickly and easily piece together an avatar and products that express me."

PicMe also received positive ratings in *satisfaction* (*MD* = 4.00, *IQR* = 1.00, Q8), *results worth effort* (*MD* = 4.00, *IQR* = 1.00, Q9), *enjoyment* (*MD* = 4.00, *IQR* = 2.00, Q10), *exploration* (*MD* = 4.00, *IQR* = 1.00, Q11), *expressivity* (*MD* = 4.00, *IQR* = 2.00, Q12), all having a median rating of 4. The participants suggested that PicMe can provide users with inspiration during personalization in terms of both content and workflow. P2 stated: "*This system provides guidance for customization in a creative way. If I was customizing something from scratch, I probably wouldn't come up with all these different scene options to pick from or consider what products I could put it on.*" P20 expressed, "*Apps that can render pictures in different art styles are common, but not ones that help you use that portrait to customize products. Following the workflow of transferring style, picking a scene, and selecting products is a creative way to make customized things.*"

6.4.3 Supporting Self-Expression. After using the personalized products for a period of time, almost all of the participants rated 4 or 5 in *meaningfulness* (*MD* = 5.00, *IQR* = 1.00, Q13), *identity* (*MD* = 4.00, *IQR* = 1.00, Q14), *self-image* (*MD* = 4.00, *IQR* = 0.25, Q15), *attention* (*MD* = 4.00, *IQR* = 2.00, Q16), and *social bond* (*MD* = 4.00, *IQR* = 1.00, Q17). The process of customization enables self-reflection and definition of the self, while the final unique product can act as an extension and ambassador of the owner's identity. P6 said, "*You can immediately recognize that this is my cap based on this portrait. The doodle of a slacker totally depicts my preference - it really captures my true essence.*" P3 added, "*I really like wearing this personalized tee. The smiley head portrait suits my playful vibe.*" P3 agreed that it is a "safe" way to express hiself, "*The cartoon portrait represents me without exposing too much personal information.*"

Our observations also indicate that personalized products have the potential to assist users in garnering positive attention, obtaining acknowledgment, and forming social bonds between themselves and others. Many participants agreed that "*I like the attention I get when people see a product with my portrait printed on it.* " Some mentioned that during the use of the product, friends, parents, or peers would think the product "*looks quite like you*" and "*suits you very much*". P7 mentioned: "*Getting feedback from my friends that the portrait actually resembles me gives me such a thrill and sense of accomplishment. Those kinds of comments reinforce that personalizing something with this avatar image was a really savvy idea.*" Also, interpersonal understanding can be enhanced through the insight of users' personalized products. As P14 described, when sharing the story of obtaining her personalized product, affirmative feedback from friends that "*Whoa, this portrait is your spitting image!*" made her feel closer and more connected to the product, and it also made her friends get to know her from a new perspective.

7 DISCUSSION

In this section, we discuss implications that have arisen through the course of designing and evaluating PicMe, as well as limitations in our current work.

7.1 Visualization of Non-abstract Personal Data

Our work focuses on the visualization of non-abstract personal data, that is, photographs, and explores its impact on product personalization. Specifically, we utilized figurative visualizations of face data, which can be easily understood and widely accepted even by people without expertise in data visualization when compared to visualizations of abstract data using charts and plots. In the user study, the participants also suggested various possibilities for figurative visualizations, such as adding data that can not be reflected solely through face data to head portraits. For example, P9 suggested, "*Would it be possible to add patterns that show my hobbies? Like if I'm into cats or reading, could I have elements related to that in the background scene?*" In our future work, we plan to explore more non-abstract data sources (e.g., videos and photographs) and visualization genres (e.g., sculptures and installations) regarding figurative visualization.

Also, we observed the potential for affective design in figurative visualization [32, 69]. Affective visualization is able to relate to, arise from, or influence emotion, which has been applied to various fields such as health and well-being as well as news and media [31]. During the process of generating personalized head portraits, we found that the participants preferred designs that can convey their emotions and feelings. For example, the participants preferred hairstyles, beards, glasses, and other elements that were similar to their visual appearance. However, when selecting components related to facial expressions, more participants chose components according to their feelings rather than the facial expression presented in the photograph. As P16 said, "*I selected a monster-like expression as I'm a goofy prankster on the inside! It represents the naughty monster side of my personality.*" In addition, many participants implied that when choosing a scene, they consider the contextual usage of the product, and subsequently select components based on the emotions they wish to communicate within that context. Examples include "*full of inspiration when working*" (P12) and "*feeling relaxed when listening to music*" (P3).

7.2 Exploring Visual Representations of Self-Image for Self-Expression

As a visual carrier of self-expression, head portraits provide the possibility to visually articulate subjective inner experiences such as self-identity, inner world, and memories [49]. Such a medium can transcend language's limitations, offering a more intuitive avenue for exploring and constructing self-identity. When using PicMe, the act of selecting, modifying, and beautifying head portraits reflects users' aspirational self-perception. In the user study, we also found that when generating portraits, the participants sometimes re-selected elements that were not perfectly aligned with their face photos (e.g., funny expressions (P16) and the slacking off scene (P12)), signaling a desire to convey an idealized self-image or a "wannabe" image to others. Also, the self-expression manifest in head portraits exerts an ongoing influence on users' self-perception when they use personalized products in their daily lives.

The participants also suggested adding more presentation modalities and content for head portraits. For example, future work can explore enabling more diverse and long-lasting avenues for self-expression. Possible research directions include expanding the representation of self-image from the head to the whole body, or adding animation and interactivity to head portraits [30, 67, 68]. P13 reflected that "*It'd be cool to make this portrait move and show off some of my signature moves.*" The application of stylized head portraits to virtual communities is also an option. As P8 said in the interview, "*I made this portrait my avatar on my gaming profile too. It lets me put my style out there for my squad to see* " Encouraging people to explore and craft visual representations of their self-image can facilitate creative self-expression, which allows individuals to better understand and communicate their identity and ideals.

7.3 AI-Supported Design for Product Personalization

According to both the quantitative and qualitative feedback, we found that the participants were satisfied with the expressiveness of the head portraits generated by PicMe. First, AI-supported design greatly improves the efficiency of creating head portraits. "*Uploading a photo, generating your head portrait, and picking the product took only a few minutes. It's awesome how fast you can go from a blank slate to having a personalized product that's exactly your style*" (P11). Second, AI-supported design is a way to help users get inspiration and collect ideas. For users with no design experience, AI-supported design greatly lowers the design barrier. "*It's great because it makes customization way more accessible. It connects personalized products, which can feel out of reach, to our daily lives. So many folks who wouldn't get to try customization can now do it in an affordable way that works for them*" (P5). However, AI-supported design still has limitations regarding head portrait generation. While AI-supported design [66] can help generate head portraits that are similar to the source face photo, some users prefer results expressing more emotions and personal characteristics. Also, the algorithm of AI-supported design depends largely on data, necessitating the construction of new models for specific artistic styles. Thus, design collaboration between AI and humans can better improve the efficiency and effectiveness of AI-supported design.

7.4 Limitations and Future Work

Our current dataset for PicMe was designed to support generating head portraits of a cartoonish and minimalist style. Our future work includes expanding the dataset such as increasing the number of components by category and adding more categories. Our algorithm requires training on specific datasets when conducting different generation tasks, which can be time-consuming and resource-intensive. Thus, our future work can explore transfer learning [84] to improve the capability of our algorithm for quickly adjusting learning strategies across datasets of different artistic styles and for effective knowledge transfer between different generation tasks. Another promising avenue for future research is investigating how to optimize system capabilities to satisfy users' requirements for personalization while retaining AI's capacity for diverse content generation, thus promoting human-AI co-creation. For example, subsequent to a user choosing a head portrait and a product, the system could suggest the location and scale for pattern prints on the product by analyzing the user's preference and the choices of other users.

8 CONCLUSION

In this work, we presented PicMe, a design support tool that generates stylized head portraits that can be used to personalize products to enable self-expression. Our experiments and user study demonstrated PicMe's effectiveness in creating stylized portraits and augmenting self-expression. Specifically, users responded positively to the portraits' quality, visual style, and ability to reflect individual identity and personal characteristics. Products incorporating the portraits were perceived as more personally relevant and as creating a deeper social bond. Future work will dive deeper into emerging applications of visualizing non-abstract personal data in identity expression and providing insights into user perceptions of data tangibility. Also, we hope this work can shed light on the design of AI techniques that actively collaborate with human designers in creating personalized portraits tailored to users' individual preferences and needs for self-expression.

ACKNOWLEDGMENTS

Nan Cao is the corresponding author. This work was supported by NSFC 62072338.

REFERENCES

[1] Bill Albert and Tom Tullis. 2022. *Measuring the User Experience: Collecting, Analyzing, and Presenting UX Metrics.* Morgan Kaufmann, Cambridge, MA, USA.
[2] Jonah Berger and Chip Heath. 2007. Where consumers diverge from others: Identity signaling and product domains. *Journal of consumer research* 34, 2 (2007), 121–134.
[3] Fadi Botros, Charles Perin, Bon Adriel Aseniero, and Sheelagh Carpendale. 2016. Go and grow: Mapping personal data to a living plant. In *Proceedings of the International Working Conference on Advanced Visual Interfaces.* ACM, New York, NY, USA, 112–119.
[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
[5] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* American Psychological Association, Washington, DC, USA.
[6] Kaidi Cao, Jing Liao, and Lu Yuan. 2018. CariGANs: Unpaired Photo-to-Caricature Translation. *ACM Trans. Graph.* 37, 6, Article 244 (2018), 14 pages.
[7] Hong Chen, Nan-Ning Zheng, Lin Liang, Yan Li, Ying-Qing Xu, and Heung-Yeung Shum. 2002. PicToon: a personalized image-based cartoon system. In *Proceedings of the tenth ACM international conference on Multimedia.* ACM, New York, NY, USA, 171–178.
[8] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25.
[9] Min Jin Chong and David Forsyth. 2022. Jojogan: One shot face stylization. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVI.* Springer, Springer, Cham, Switzerland, 128–152.
[10] Keisha M Cutright, Tülin Erdem, Gavan J Fitzsimons, and Ron Shachar. 2014. Finding brands and losing your religion? *Journal of Experimental Psychology: General* 143, 6 (2014), 2209.
[11] Audrey Desjardins and Timea Tihanyi. 2019. ListeningCups: A case of data tactility and data stories. In *Proceedings of the 2019 on designing interactive systems conference.* ACM, New York, NY, USA, 147–160.
[12] Matthijs Douze, Arnau Ramisa, and Cordelia Schmid. 2011. Combining attributes and fisher vectors for efficient image retrieval. In *CVPR 2011.* IEEE, IEEE, 745–752.
[13] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and mind: a study of avatar personalization in three virtual worlds. In *Proceedings of the SIGCHI conference on human factors in computing systems.* ACM, New York, NY, USA, 1151–1160.
[14] Julia Gong, Yannick Hold-Geoffroy, and Jingwan Lu. 2020. Autotoon: Automatic geometric warping for face cartoon generation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision.* IEEE, Piscataway, NJ, USA, 360–369.
[15] Mar Gonzalez-Franco and Tabitha C Peck. 2018. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI* 5 (2018), 74.
[16] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (2020), 139–144.
[17] Pascalle CM Govers and Ruth Mugge. 2004. I love my Jeep, because its tough like me: The effect of product-personality congruence on product attachment. In *Proceedings of the fourth international conference on design and emotion, Ankara, Turkey*, Vol. 15. Taylor & Francis, London, UK, 1–14.
[18] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* IEEE, Piscataway, NJ, USA, 1125–1134.
[19] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* IEEE, Piscataway, NJ, USA, 1125–1134.
[20] Wonjong Jang, Gwangjin Ju, Yucheol Jung, Jiaolong Yang, Xin Tong, and Seungyong Lee. 2021. StyleCariGAN: caricature generation via StyleGAN feature map modulation. *ACM Transactions on Graphics (TOG)* 40, 4 (2021), 1–16.
[21] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2017. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196* (2017).
[22] Rohit Ashok Khot, Larissa Hjorth, and Florian'Floyd' Mueller. 2014. Understanding physical activity through 3D printed material artifacts. In *Proceedings of the SIGCHI conference on human factors in computing systems.* ACM, New York, NY, USA, 3835–3844.
[23] Rohit Ashok Khot, Jeewon Lee, Deepti Aggarwal, Larissa Hjorth, and Florian'Floyd' Mueller. 2015. Tastybeats: Designing palatable representations of physical activity. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 2933–2942.
[24] Heejung S Kim and Deborah Ko. 2011. Culture and self-expression. In *The self.* Psychology Press, 325–342.
[25] Junho Kim, Minjae Kim, Hyeonwoo Kang, and Kwanghee Lee. 2019. U-gat-it: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation. *arXiv preprint arXiv:1907.10830* (2019).
[26] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*, Yoshua Bengio and Yann LeCun (Eds.). OpenReview.net.
[27] Rune Klevjer. 2006. *What is the avatar? Fiction and embodiment in avatar-based singleplayer computer games.* Ph. D. Dissertation. The University of Bergen, Bergen, Norway.
[28] Ming-Hsun Kuo, Yong-Liang Yang, and Hung-Kuo Chu. 2016. Feature-Aware Pixel Art Animation. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, Wiley Online Library, 411–420.
[29] Stacey Kuznetsov and Eric Paulos. 2010. Rise of the expert amateur: DIY projects, communities, and cultures. In *Proceedings of the 6th Nordic conference on human-computer interaction: extending boundaries.* ACM, New York, NY, USA, 295–304.
[30] Xingyu Lan, Yang Shi, Yanqiu Wu, Xiaohan Jiao, and Nan Cao. 2021. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 933–943.
[31] Xingyu Lan, Yanqiu Wu, and Nan Cao. 2024. Affective Visualization Design: Leveraging the Emotional Impact of Data. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 1–11.
[32] Xingyu Lan, Yanqiu Wu, Yang Shi, Qing Chen, and Nan Cao. 2022. Negative emotions, positive outcomes? exploring the communication of negativity in serious data stories. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.* ACM, 1–14.
[33] Micah Lanier. 2020. Avatar Illustration System. https://www.figma.com/community/file/829741575478342595. Accessed: 2023.
[34] Nguyen Kim Hai Le, Yong Peng Why, and Golam Ashraf. 2011. Shape stylized face caricatures. In *International Conference on Multimedia Modeling.* Springer, Cham, Switzerland, 536–547.
[35] Moon-Hwan Lee, Oosung Son, and Tek-Jin Nam. 2016. Patina-inspired personalization: personalizing products with traces of daily use. In *Proceedings of the 2016 ACM conference on designing interactive systems.* ACM, New York, NY, USA,

251–263.
[36] Shaohui Lei, Xianqing Wang, Leiqing Peng, and Yulang Guo. 2021. "I" seek differentiation and "we" seek assimilation: the impact of self-expressive customization on consumers' willingness to pay a premium. *Journal of Product & Brand Management* 30, 5 (2021), 691–706.
[37] LeoKu. 2021. vue-color-avatar. https://github.com/Codennnn/vue-color-avatar. Accessed: 2023.
[38] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. 2022. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *International Conference on Machine Learning*. PMLR, San Diego, CA, USA, 12888–12900.
[39] Jingyi Li, Wilmot Li, Sean Follmer, and Maneesh Agrawala. 2021. Automated Accessory Rigs for Layered 2D Character Illustrations. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. ACM, 1100–1108.
[40] Tzu-Mao Li, Michal Lukáč, Gharbi Michaël, and Jonathan Ragan-Kelley. 2020. Differentiable Vector Graphics Rasterization for Editing and Learning. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 39, 6 (2020), 193:1–193:15.
[41] Yan Li, Ruiping Wang, Haomiao Liu, Huajie Jiang, Shiguang Shan, and Xilin Chen. 2015. Two birds, one stone: Jointly learning binary code for large-scale face image retrieval and attributes prediction. In *Proceedings of the IEEE International Conference on Computer Vision*. IEEE, 3819–3827.
[42] Fangpen Lin. 2021. Avataaars generator. https://github.com/fangpenlin/avataaars-generator. Accessed: 2023.
[43] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep Learning Face Attributes in the Wild. In *Proceedings of International Conference on Computer Vision (ICCV)*. IEEE, Piscataway, NJ, USA, 3730–3738.
[44] Arnold M Lund. 2001. Measuring usability with the use questionnaire12. *Usability interface* 8, 2 (2001), 3–6.
[45] Sampada Marathe and S Shyam Sundar. 2011. What drives customization? Control or identity?. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 781–790.
[46] Paul R Messinger, Xin Ge, Eleni Stroulia, Kelly Lyons, Kristen Smirnov, and Michael Bone. 2008. On the relationship between my avatar and myself. *Journal For Virtual Worlds Research* 1, 2 (2008), 1–17.
[47] Midjourney. 2022. Midjourney. https://www.midjourney.com/home/.
[48] Kiet Tuan Nguyen, Phuong Van Nguyen, Khoa Anh Ngoc Le, and Thao Danh Doan Bui. 2018. Identity Representation in Customization. A Case of Nike Shoes. *Int. J Sup. Chain. Mgt Vol* 7, 5 (2018), 582.
[49] Jie Nie, Peng Cui, Yan Yan, Lei Huang, Zhen Li, and Zhiqiang Wei. 2014. How your portrait impresses people? Inferring personality impressions from portrait contents. In *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, New York, NY, USA, 905–908.
[50] Bettina Nissen, John Bowers, Peter C Wright, Jonathan David Hook, and Christopher Newell. 2014. Volvelles, domes and wristbands: embedding digital fabrication within a visitor's trajectory of engagement. In *Proceedings of the 2014 conference on Designing interactive systems*. Association for Computing Machinery (ACM), ACM, New York, NY, USA, 825–834.
[51] OpenAI. 2021. DALL·E: Creating Images from Text - OpenAI. https://openai.com/blog/dall-e/.
[52] Daphna Oyserman. 2009. Identity-based motivation and consumer behavior. *Journal of Consumer Psychology* 19, 3 (2009), 276–279.
[53] Or Patashnik, Zongze Wu, Eli Shechtman, Daniel Cohen-Or, and Dani Lischinski. 2021. Styleclip: Text-driven manipulation of stylegan imagery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, Piscataway, NJ, USA, 2085–2094.
[54] Brett W Pelham. 1993. The idiographic nature of human personality: examples of the idiographic self-concept. *Journal of Personality and Social Psychology* 64, 4 (1993), 665.
[55] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*. PMLR, San Diego, CA, USA, 8748–8763.
[56] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PMLR, San Diego, CA, USA, 8748–8763.
[57] Americus Reed II, Mark R Forehand, Stefano Puntoni, and Luk Warlop. 2012. Identity-based consumer behavior. *International Journal of Research in Marketing* 29, 4 (2012), 310–321.
[58] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, Cham, Switzerland, 234–241.
[59] Daniel Roth, Jean-Luc Lugrin, Dmitri Galakhov, Arvid Hofmann, Gary Bente, Marc Erich Latoschik, and Arnulph Fuhrmann. 2016. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE Virtual Reality (VR)*. IEEE, Piscataway, NJ, USA, 277–278.
[60] Christina Saenger, Veronica L Thomas, and Jennifer Wiggins Johnson. 2013. Consumption-focused self-expression word of mouth: A new scale and its role in consumer research. *Psychology & Marketing* 30, 11 (2013), 959–970.
[61] Akash Saravanan and Matthew Guzdial. 2022. Pixel vq-vaes for improved pixel art representation. *arXiv preprint arXiv:2203.12130* (2022).
[62] Kim Sauvé, Saskia Bakker, Nicolai Marquardt, and Steven Houben. 2020. LOOP: Exploring physicalization of activity tracking data. In *Proceedings of the 11th nordic conference on human-computer interaction: shaping experiences, shaping society*. ACM, New York, NY, USA, 1–12.
[63] Yang Shi, Nan Cao, Xiaojuan Ma, Siji Chen, and Pei Liu. 2020. EmoG: Supporting the Sketching of Emotional Expressions for Storyboarding. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, ACM, 393:1–393:12.
[64] Yang Shi, Siji Chen, Pei Liu, Jiang Long, and Nan Cao. 2022. Colorcook: Augmenting color design for dashboarding with domain-associated palettes. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–25.
[65] Yichun Shi, Debayan Deb, and Anil K Jain. 2019. Warpgan: Automatic caricature generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Piscataway, NJ, USA, 10762–10771.
[66] Yang Shi, Tian Gao, Xiaohan Jiao, and Nan Cao. 2023. Understanding design collaboration between designers and artificial intelligence: A systematic literature review. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–35.
[67] Yang Shi, Xingyu Lan, Jingwen Li, Zhaorui Li, and Nan Cao. 2021. Communicating with motion: A design space for animated visual narratives in data videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13.
[68] Yang Shi, Zhaorui Li, Lingfei Xu, and Nan Cao. 2021. Understanding the design space for animated narratives applied to illustrations. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 1–6.
[69] Yang Shi, Pei Liu, Siji Chen, Mengdi Sun, and Nan Cao. 2022. Supporting expressive and faithful pictorial visualization design with visual style transfer. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 236–246.
[70] Pablo Stanley. 2010. Open Peeps. https://www.openpeeps.com. Accessed: 2023.
[71] Pablo Stanley. 2020. Avataaars Dataset. https://www.avataaars.com/. Accessed: 2023.
[72] Carmen Stavrositu and S Shyam Sundar. 2008. Can blogs empower women? Designing agency-enhancing and community-building interfaces. In *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2781–2786.
[73] Kil-Soo Suh, Hongki Kim, and Eung Kyo Suh. 2011. What if your avatar looks like you? Dual-congruity perspectives for avatar use. *MIs Quarterly* 35, 3 (2011), 711–729.
[74] S Shyam Sundar, Jeeyun Oh, Saraswathi Bellur, Haiyan Jia, and Hyang-Sook Kim. 2012. Interactivity as self-expression: A field experiment with customization and blogging. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 395–404.
[75] Sabine Trepte and Leonard Reinecke. 2010. Avatar creation and video game enjoyment: Effects of life-satisfaction, game competitiveness, and identification with the avatar. *Journal of Media Psychology: Theories, Methods, and Applications* 22, 4 (2010), 171.
[76] Andrew Turpin and Falk Scholer. 2006. User performance versus precision measures for simple search tasks. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, New York, NY, USA, 11–18.
[77] Yael Vinker, Yuval Alaluf, Daniel Cohen-Or, and Ariel Shamir. 2023. Clipascene: Scene sketching with different types and levels of abstraction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, 4146–4156.
[78] Yael Vinker, Ehsan Pajouheshgar, Jessica Y Bo, Roman Christian Bachmann, Amit Haim Bermano, Daniel Cohen-Or, Amir Zamir, and Ariel Shamir. 2022. Clipasso: Semantically-aware object sketching. *ACM Transactions on Graphics (TOG)* 41, 4 (2022), 1–11.
[79] James Whiting and Kevin Hannam. 2015. Creativity, self-expression and leisure. *Leisure studies* 34, 3 (2015), 372–384.
[80] Zhihao Yao, Qirui Sun, Beituo Liu, Yao Lu, Guanhong Liu, and Haipeng Mi. 2023. InkBrush: A Flexible and Controllable Authoring Tool for 3D Ink Painting. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, 1–3.
[81] Ruimao Zhang, Liang Lin, Rui Zhang, Wangmeng Zuo, and Lei Zhang. 2015. Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification. *IEEE Transactions on Image Processing* 24, 12 (2015), 4766–4779.
[82] Yang Zhao, Diya Ren, Yuan Chen, Wei Jia, Ronggang Wang, and Xiaoping Liu. 2022. Cartoon Image Processing: A Survey. *International Journal of Computer Vision* 130, 11 (2022), 2733–2769.

[83] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*. IEEE, Piscataway, NJ, USA, 2223–2232.

[84] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. 2020. A comprehensive survey on transfer learning. *Proc. IEEE* 109, 1 (2020), 43–76.