Capturing Subjective Aesthetic Expertise through Designer-Like Learning and Practice

Yaliang Chuang* **and Poyang David Huang**

Department of Industrial Design, Eindhoven University of Technology
De Zaale, 5612AJ, Eindhoven, the Netherlands
y.chuang@tue.nl

Abstract

Aesthetics is a crucial aspect of design that plays a critical role in the creation process and customers' perception of outcomes. However, aesthetic expressions are highly subjective and nuanced. It often relies on designers' experiences as well as many trials and errors to get it right. This research first investigated how designers and artists curated aesthetic materials and utilized them in their daily practice. Then, we developed an interactive Style Agent system to collect designers' aesthetic perceptions of visual materials. We used typographic posters as examples in this research. To evaluate the performance, we conducted a user study with six designer participants divided into two groups. Group 1's participants helped use our interactive interface to rate 156 poster stimuli's according to their perceptions on 30 adjectives separately. By processing the data using Concept Activation Vectors (CAV, a machine learning program), the system generates models to predict the aesthetic qualities of other one thousand new design examples. Our prototype system also provides a three-dimentionsal visualization to assist designers see the design space and every participant was asked to picked inspirational examples to design a poster for a local gallery in 20 minutes. We collect quantitative and qualitative data to assess the system's performance. Firstly, we investigated the correlations between Group 1 participants' ratings and the CAV's outcomes. All of the 90 data pairs showed significant correlations [$r(154)$ is higher than .711 (Median = .822; SD = .048), $p < .05$]. Furthermore, after completing the poster design assignment, all participants expressed that the design examples and 3D visualization potentially boosted their creativity in their creation process. The average of Creativity Support Index is 73.9. Overall, our study shows that a small number of data could achieve good performance through the interactive teaching and visualization system developed with human-like learning characteristics (Langley 2022).

Background

Aesthetics is a esential aspect of design that plays a critical role in people's perception and experience of products, services, and environments. In practice, designers often spend significant time experimenting with many ideas to create aesthetic qualities that match the clients' or target customers' tastes. However, it often relies on designers' experiences, tacit knowledge, and many trials and errors. For instance, when Tyler Hobbs created the Haecceity series of generative artworks, he found it very challenging to examine the 950 images generated by the algorithm developed by himself (Hobbs 2014). He looked at each of them and studied the best images' compositional strength, balance, rhythm, and quality of detail. After spending a significant amount of time with trial-and-error examination, he narrowed the images down to 24 and chose 7 of them, which complement each other, show the range of the program, and generally work as a series.

Another pioneer furniture design example is Schmitt and Weib's chair project which experimented with a human-AI co-design system to create innovative armchairs based on the aesthetics of iconic designs (Schmitt 2019), which. They applied the DCGAN (deep convolutional generative adversarial networks) machine learning method to train a model with an original dataset consisting of 600 images of iconic 20th-century chairs. The AI system captured some features and then generated a lot of rough sketches (see the most left side images in Fig.1). Those blurry images stimulated designers' imaginations and 'pushed them away from usual threads of

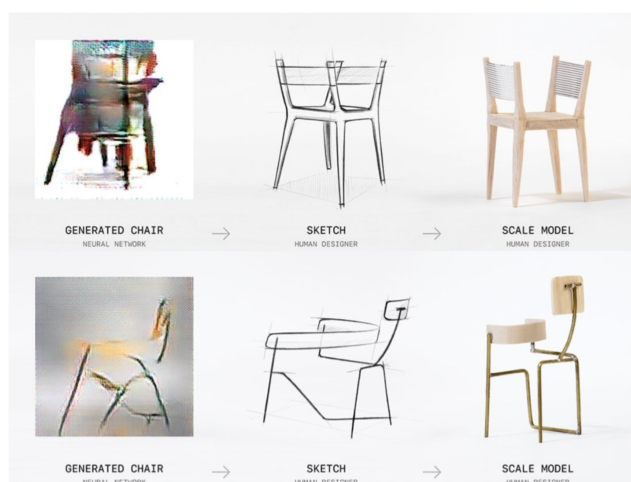

Figure 1: The design and fabrication process of transforming the AI-generated images into a sketch and physical model. (Schmitt and Weiß 2018).

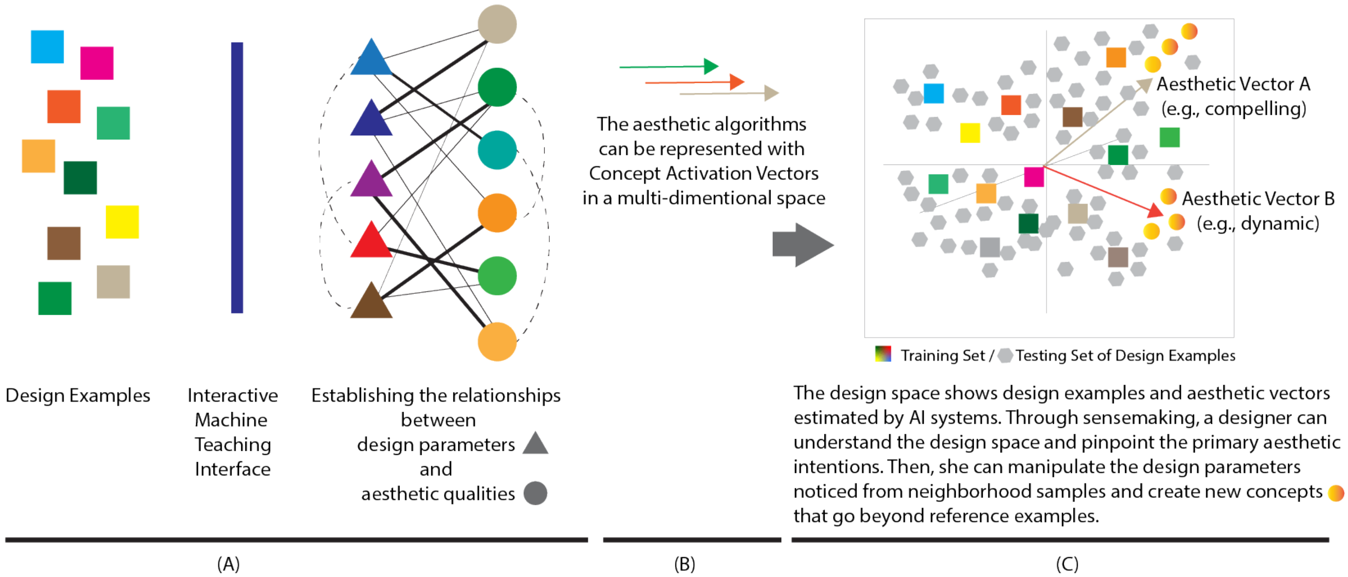

Figure 2: The framework of this research project: (A) Developing an interactive machine teaching interface for engaging designers to teach AI systems with design examples. (B) Using machine learning techniques [e.g., Concept Activation Vectors (Kim et al. 2018)] to transform aesthetic algorithms into computational forms. (C) Visualizing the design space with examples and aesthetic vectors generated from B. Designers can not only use the map to identify new opportunities but also manipulate the noticed design parameters (such as shapes, colors, or materials) to experiment with creative designs for conveying certain aesthetic qualities (e.g., expressing energetic and dynamic feelings with balanced quality).

thinking towards unusual ideas they might not have had otherwise" (p. 2). Designers reviewed hundreds of AI-generated images, selected four potential ones, and further transformed them into design sketches and physical models with their craftmanship expertise (Fig.1). Although recent generative AI applications (e.g., Dall·E or MidJourney) show an impressive improvement in the outcomes' quality, the dominant prompt engieenring is not intuitive for designers to steer the ideation process with particular aesthetic visions or intentions. It was also found that AI generated contents shared obvious similarities and lacking the diversity comparing human generated outcomes (Dell'Acqua et al. 2023). Perhaps this is because the existing AI systems lack knowledge of aesthetics and subjective feelings. As a result, designers still need to examine vast amounts of generated outcomes and adjust their inputs by trial and error. **What if the designers can teach AI what aesthetic qualities entail with examples they collected or created?** We applied Langley's frameworks of human-like learning approach (Langley 2022) and selected three of the seven characteristics to develop an Style Agent prototype system. We utilize designer's prior experiences and repository of design examples to teach systems with their aesthetic styles in a piecemeal manner. The interactive teaching interface was developed with a modular structure based on Kansei Engieering (Lévy and Yamanaka 2009) methodology. Powered by interactive machine-teaching techniques (van der Stappen and Funk 2021), the system can learn the essence of designers' aesthetic expertises and build the user's model as a computational representation of their design style (see Fig.2B). For instance, the Concept Aviation Vectors (CAV) method (Kim et al. 2018) can generate mathematical functions to represent the aesthetic qualities in multidimensional design space. Furthermore, we use the user's CAV model to predict the aesthetic perceptions of new design examples (see Fig. 2C). This design space can help designers to define a concrete direction for a given assignment and examine the design references to examine critical design parameters and how they might affect particular aesthetic qualities. In this study, we used the poster design as a medium to investigate its application and efficacy.

Methods

To help reviewers and readers understand the research concept presented in this proposal and evaluate its novelty and feasibility, we use poster design as an example to illustrate how we envision the system could be developed to investigate the research questions described in the Background section. In this chapter, we will first explain the system process of collecting design examples and training the machine with design participants. Then, we will describe how we use Google AI (Google 2022) and Mood Board Search (Research and Projects 2022) to analyze the training data and build designer's style models to estimate the aesthetic qualities of other one thousand design examples. Thirdly, we will present the data visualization interface for displaying the machine-learning results to help designers browse the examples in a multi-dimensional space and curate relevant examples to develop their own concepts. We hypothesize that designers might not fully agree with the results, yet the dis-

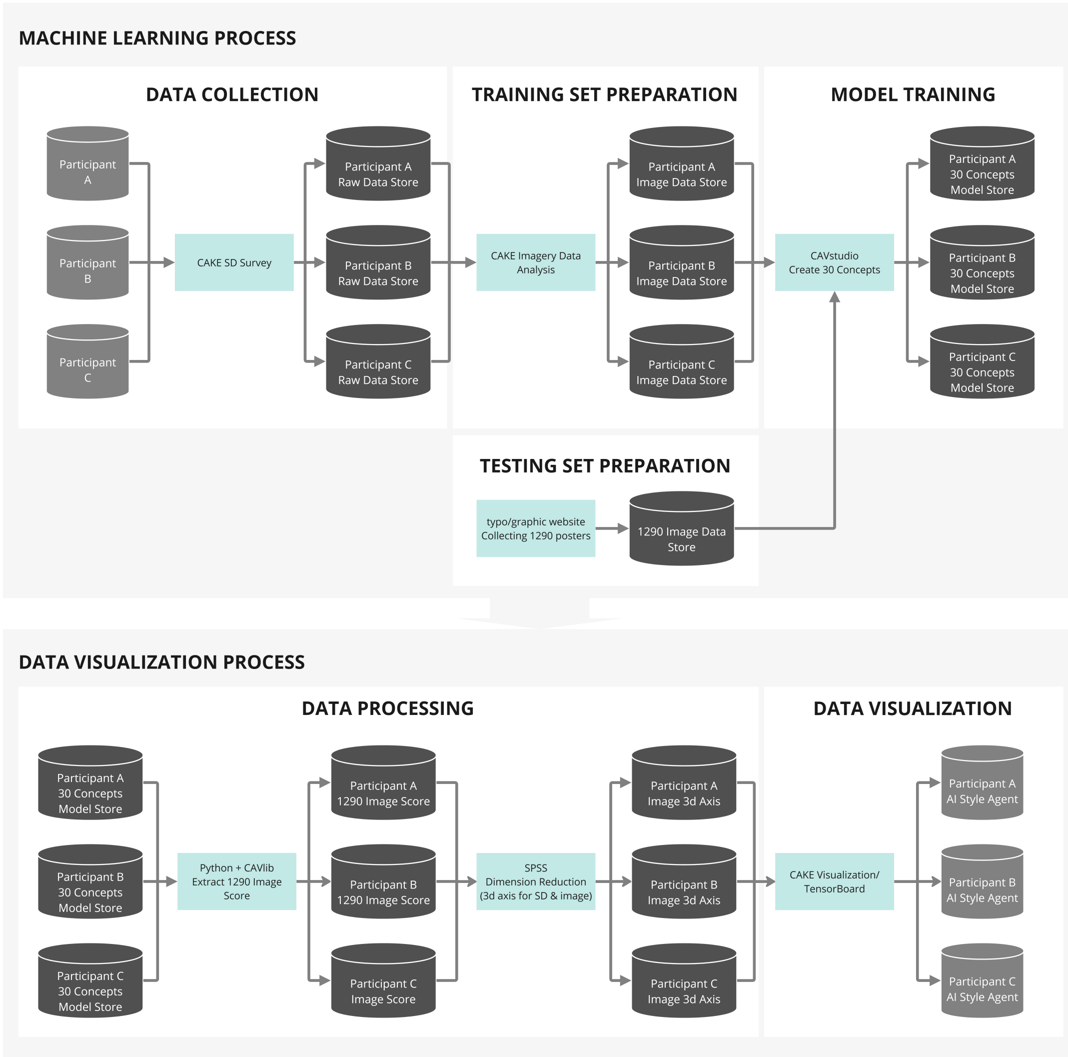

Figure 3: The overview process of the proposed Style Agent system. It includes two main phases: machining learning and visualization. The primary components of each one were presented in this diagram.

crepancy could trigger designers to think about why particular designs were perceived differently by other users or the AI. This could inspire designers to define a more suitable language to convey their expressivities and ensure their designs could be interpreted as they intended by the audience or consumers. At the end of this section, we will describe a user study conducted with six designer participants and the preliminary findings.

The Overall Process of Designing the Style Agent

This study focuses on the influence of a designer's style on design activities. In order to observe the extent of visualization of a designer's style and its influence on the design process, this study can be divided into two parts: the machine learning phase and the data visualization phase. Fig.3) presents the overall process and the primary components included in each phase.

Phase 1: Machine Learning Stage

This stage includes four parts. Firstly, defining the data structure with aesthetic vocabularies. We will use the semantic differential method commonly utilized in emotional design or Kansei engineering research (Lévy and Yamanaka 2009). Secondly, we collect high-quality examples from design platforms. Thirdly, we use the CAKE system (Chuang and Chen 2007) to collect participants' perceptions as a way to train the data. Lastly, we use Google's AI and Mood Board search (Google 2022) to process the data and build machine learning models for each of the aesthetic scales.

Defining the Data Structure and Aesthetic Vocabularies Before moving into the design process, we first analyzed the semantic terms used in studies on graphic design (Zhao, Cao, and Lau 2018) and product shape (Hsiao and Chen 2006) and function (Sevener 2003; Chuang, Chen, and Chuang 2008) to find adjectives that could be used to de-

Total evaluation (13)	Rational - Emotional	Avant-garde - Conservative	Consistent – Inconsistent	Dislike - Like	Elegant - Not elegant
	Boring - Amazing	Modest - Exaggerated	Ordinary - Extraordinary	Unattractive - Attractive	
	Uncreative - Creative	Exciting - Calm	Complex - Simple	Bad-looking - Beautiful	
Image (9)	Futuristic - Nostalgic	Cute - Not cute	Balance - Unbalance	Humor - Serious	Dynamic - Static
	Comfortable - Uncomfortable	Lifeless - Energetic	Formal - Casual	Feminine - Masculine	
Perception (6)	Varied - Monotonous	Refined - Rough	Streamlined - Rugged		
	Cold - Warm	Light - Heavy	immature - Mature		
Shape element (2)	Geometric - Organic	Symmetrical - Asymmetrical			

Figure 4: The chosen 30 semantic differential scales used in the presented study.

scribe the aesthetic qualities of two- or three-dimensional designs. A set of opposite adjectives can form a semantic differential (SD) pair. Then, we collaborated with another researcher to select and define 30 sets of SD (Fig.4) for the evaluation of visual materials through informal evaluation, referring to the hierarchy of the Kansei evaluation model established by Kobayashi et al. (Kobayashi, Kinumura, and Higashi 2016).

Collecting High-quality Design Examples The visual materials to be evaluated were 156 high-quality posters randomly selected from the typo/graphic posters website (typographicposters.com) because the posters contained many elements that could be used as design styles, such as color, shape, and typography, and were easy to describe the feeling in words.

Training the Data with Design Participants The survey interface contains the poster to be evaluated and the SDs (Fig.5). To confirm that the semantic data of the posters could be stored completely, the researchers personally tested it to collect 30 sets of SD for 156 posters. We invited another 3 participants who majored in Industrial Design to perform the same collection task, with the aim of using these data as a training set for creating a personal style model for each participant. In average, it took 4.5 hours for a participant to complete all the training tasks.

Using AI to Analyze the Data and Make Predictions After collecting all the data, we used CAKE's imagery data analysis system to read each participant's XML file. The system generated a positive and a negative correlation folder for each adjective according to the poster's score on different adjectives. CAKE system helps us to convert each participant's semantic ratings of all posters from numbers to actual image examples, which will be used as a training set for the later machine-learning process. As for the testing set, we collected 1134 high-quality posters randomly from the same typo/graphic posters website.

To train the machine to recognize abstract adjectives and visual posters, we chose CAVstudio from Google's Mood Board Search (Google 2022) as the machine training tool. CAVstudio uses a machine learning method called concept activation vectors (CAVs) to identify visual concepts. It works by finding a direction or "vector" in a high-dimensional space that represents a particular concept (Kim

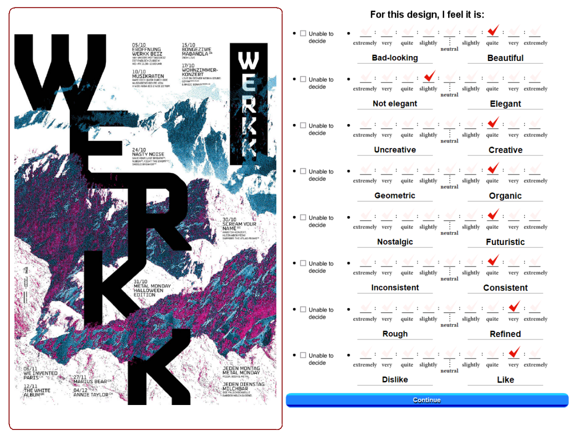

Figure 5: The semantic differential survey of Chuang's CAKE system (Chuang and Chen 2007) was used to train the data by asking design participants to rate their aesthetic perceptions of each of the 156 poster examples.

et al. 2018). The user can train the model by giving it a positive and negative image of the concept, a training approach that is in line with the concept that we store posters separately according to semantic word scores in the data collection phase.

For example, in Fig.6, the 'Extraordinary' concept is being trained, images that match this concept are placed in the 'It is...' area and images that do not match this concept are placed in the 'It is not...' area. Then, select the training model and press 'Learn Concept' to perform machine learning. Once the CAV is trained, it can be compared to the new image in the 'Concept results' area to see how similar it is by looking at the angle between the CAV and the activation of the new image, a technique known as cosine similarity. The score ranges from -1 to 1, and the higher the score, the more similar the image is to the trained concept and the correlation is expressed in the interface in terms of image size, with larger images representing higher similarity to the concept.

However, it is not intuitive to present 1290 posters (156 trained images and 1134 testing ones) in this way because the user needs to keep scrolling up and down to see other images. It is difficult to view and compare the differences

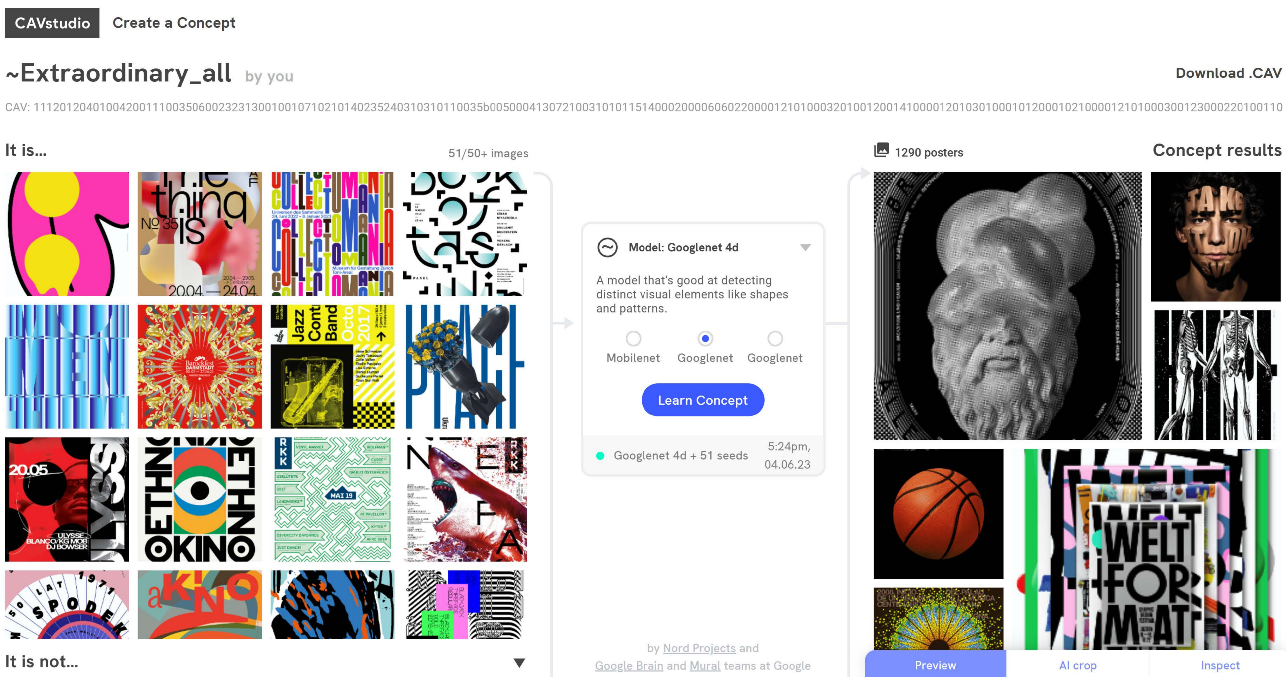

Figure 6: The example of training an ML model in CAVstudio (Research and Projects 2022) to identify images that match the user's concept of 'Extraordinary.'

between images at the same time, so we only use CAVstudio as a model training tool, not as a data visualization tool.

Data Visualization Stage

Data Processing The idea was to display the 30 adjectives in the space in the form of axes so that each poster could be distributed in the 3D space according to its score for each semantic pair. To find out the relationship between adjectives and make the analysis easier, we used the statistical technique of principal component analysis (PCA) in the factor analysis of IBM SPSS to reduce the dimensionality of the data. Eventually, each participant's 30 sets of SDs for all posters were reduced to 3 groups, which could respectively correspond to the elements composing the X, Y, and Z axes, and their scores represented their contributions to the coordinates (Fig.7 Left).

Data Visualization We used the CAKE Visualization tool to display the coordinates of stimuli from factor analysis and multi-dimensional scaling analysis, which are most commonly used in imagery-related research, in interactive virtual space (Chuang and Chen 2008). The interface displays the posters in the canvas, and users can rotate the multi-dimensional space freely to observe the relationship of different posters in the space or display only a few posters (Fig.7). The system also provides the function of temporarily saving the clicked posters and presenting them in a clearer way on the right panel so that users can click posters for observation or remove specific posters.

A Preliminary User Evaluation on the Prototype System

We referenced the experimental design of Kang et al. (Kang et al. 2021) and developed different tasks for participants to complete. There were five design participants recruited in this pilot study. In the following, we report the experimental procedure, data collection methods, and preliminary findings.

Procedure The experimental process basically consisted of four stages. In the first stage, we explained the purpose of the study to the participants and asked them to introduce to us the design concepts and process of the posters they had designed in the past year, in order to let us understand the design process and tools that the participants were accustomed to use in designing, and at the same time to let the participants recall their own design experience as a basis for subsequent comparison. In the second stage, we first showed the participants our system, explained how we used AI to design the system and how to operate the system, and then asked them to select the most appropriate adjective for XYZ axis from the given adjective pairs to represent that axis, in order to allow participants to explore and visualize how the posters in the system relate to their own styles. In the third stage, we asked participants to design a poster for an art exhibition using our system as a source of inspiration within 20 minutes. The design assignment is "to design a poster for an art exhibition about colors which will be held in a local gallery." The total duration of the experiment was approximately 1 hour. Each task was followed by a qualitative survey.

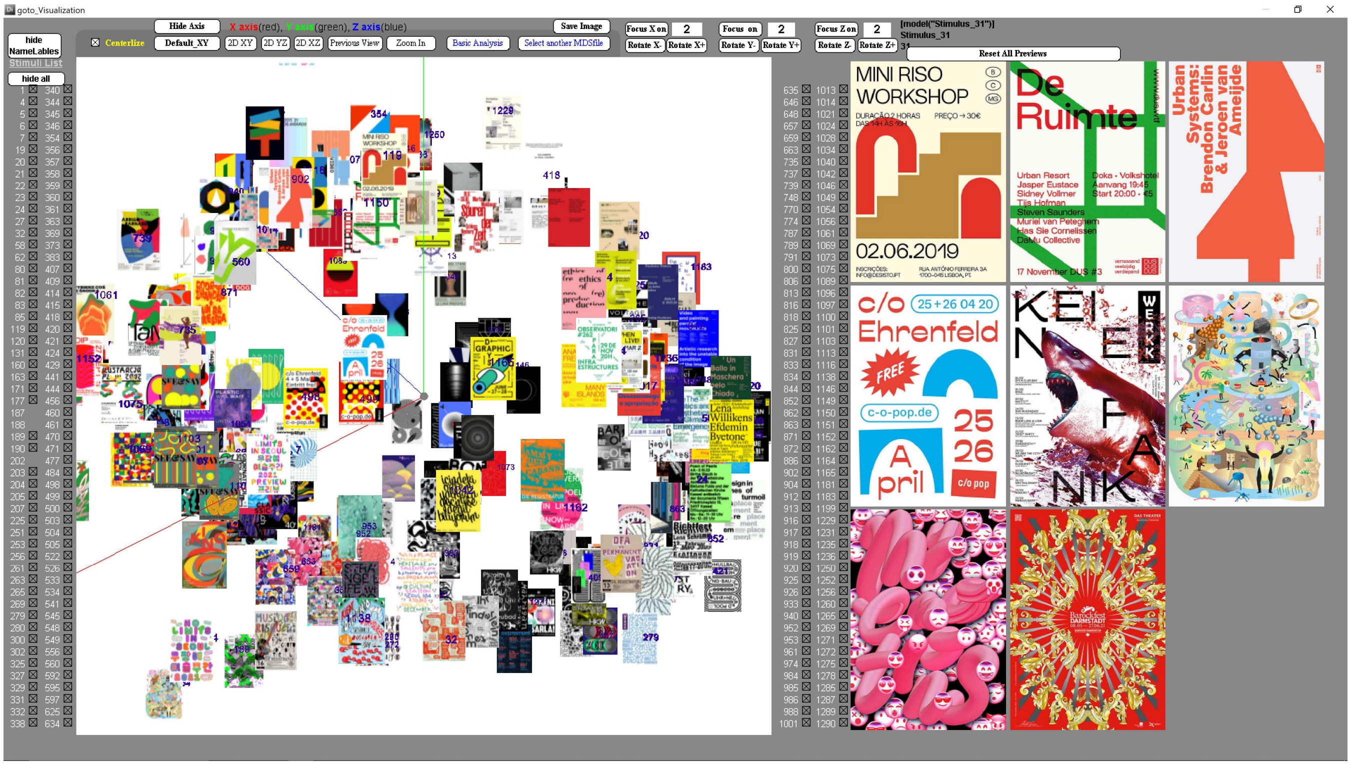

Figure 7: We used the CAKE visualization tool (Chuang and Chen 2007) to display the design space with examples and their coordinates analyzed from Google AI. A user can rotate the space to observe the distributions of designs and click the examples to preview them on the right. This can help designers to curate design examples and analyze their parameters to inspire them to create new designs that can convey particular aesthetic qualities.

Data Collection Methods To measure the invisible creative process in a more comprehensive way, we used both quantitative and qualitative research methods in evaluating the system's functionality to obtain participants' suggestions for improvement. At the end of the first task, we asked them to indicate their reasons for choosing the adjective pair and how satisfied they were with how their style was presented in the system. At the end of the second task, we asked them to describe the poster they designed, the design process, the difficulties they encountered, how the system helped them in the design process or how they would like to improve the system. Finally, we asked them to fill out a Creativity Support Index (CSI) questionnaire (Cherry and Latulipe 2014) to obtain a score for the system. CSI is a standardized psychometric tool for assessing the perceived creativity support of a tool, including collaboration, enjoyment, exploration, expressiveness, immersion, and worthiness of effort.

Preliminary Findings and Reflections To evaluate the performance of CAV's outcomes, we conducted a Pearson Correlation analysis. Among the 90 data pairs of three participants in Group 1, the results are all significant ($p < .05$) and the $r(154)$ is higher than .711 (Median = .822; SD = .048). Two example char was shown in Fig.8. The amount of data used to teach the system is between 60 to 80 items in each of the aesthetic attributes. This shows a small number of dataset could achieve good performance through the interactive teaching system developed with the human-like-learning characteristics (Langley 2022).

In order to understand the extent to which the system reflects the personal style of the designers, at the beginning of the experiment, we explained to the participants that the scores of the images they had helped to collect were used as the training set for their style agent and that the style agent predicted possible scores for other new images for them based on this information, and presented them in the space via a data visualization. This helped them to quickly understand the rationale behind the system, to recognize that the system was presented in a logical rather than random way, that it was partly personal to them, and to develop a sense of identity with the system. After getting familiar with the 3D visualization, all participants said the Style Agent captured their style preferences because the distribution of examples in the multi-dimensional space echoes the differences of stimuli's aesthetic qualities perceived by the participant.

When explaining the reasons for defining axial adjectives or the poster design process, the participants always used the preview function to select a few posters and then selected poster features as examples for ideation. This process highlights the need to examine the features of the selected posters and validates the ability of images to help convey abstract aesthetic concepts (Eckert and Stacey 2000).

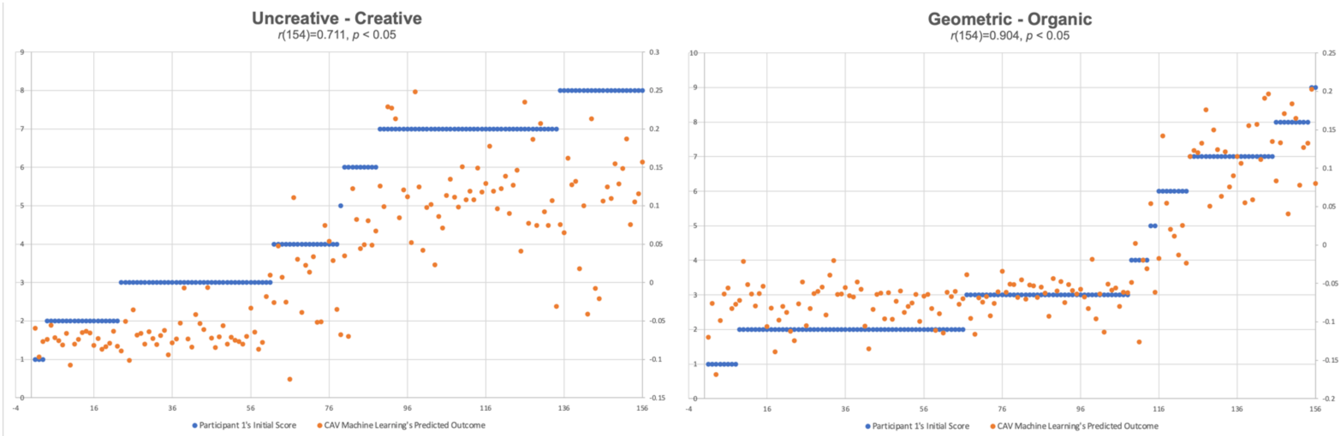

Figure 8: Those two chars were produced based on participant 1's data of the particular aesthetic values of 156 poster stimuli. The left one is the Uncreative-Creative aesthetic attribute which shows the lowest correlation between the participants' initial scores (blue dots) and the machine learning's prediction (red dots). In contrast, the right one of Geometric-Organic attribute show the highest significant correlation among all data pairs analyzed.

Figure 9: Posters designed by participants with the Style Agent system.

The results of the CSI scale for the two groups of participants revealed that both groups of participants who were personally involved in the data collection and those who were not involved in the data collection gave similar scores. The mean score for the first group was 74.2, the second group was 73.7, and the mean score for both groups was 73.9. This indicates that the prototype system can potentially boost participants' creativity in their creation process. Therefore, we are currently collaborating with professional designers or artists to investigate their process and use the insight to refine our research framework (Fig.2) and system design.

In the poster design task, we allowed participants to use our system as a design starting point without explicitly stating that they could not imitate it. The participants' designs are shown in Fig.9.

Even though our system is not like an AI image generation tool that tends to guide creativity orientation, there is still the problem of "fixation" of design ideas that often occurs in image generation AI tools, where the designer's creativity is influenced by the original reference object, resulting in overly similar design outcomes. The possible reason for this is that the designers chose to use elements from the original reference objects for their designs due to insufficient time. In the future study, we will adjust the design task in the experiment setup.

One challenge we are researching on is to acquire designers' practical knowledge and refine the system's model rapidly through observing how designers using the reference exemplars in their creation process, i.e., designing a new poster. This aligns to Langley's human-like learning characteristics and we belive the interactive machine teaching approach could foster incremental learning and refinement. We are looking forward to present our project and the latest progress in this AAAI-24 Spring Symposium and to learn from colleagues and experts.

References

Cherry, E.; and Latulipe, C. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction*, 21(4): 21:1–21:25.

Chuang, Y.; and Chen, L.-L. 2007. CAKE: An Extensible

Kansei Survey and Analysis System. *Journal of Design*, 12(3): 63–80.

Chuang, Y.; and Chen, L.-L. 2008. How to rate 100 visual stimuli efficiently. *International Journal of Design*, 2(1): 31–43.

Chuang, Y.; Chen, L.-L.; and Chuang, M.-C. 2008. Computer-based rating method for evaluating multiple visual stimuli on multiple scales. *Computers in Human Behavior*, 24(5): 1929–1946.

Dell'Acqua, F.; McFowland, E.; Mollick, E. R.; Lifshitz-Assaf, H.; Kellogg, K.; Rajendran, S.; Krayer, L.; Candelon, F.; and Lakhani, K. R. 2023. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. BoI Conference.

Eckert, C.; and Stacey, M. 2000. Sources of inspiration: a language of design. *Design Studies*, 21(5): 523–538.

Google. 2022. Mood Board Search.

Hobbs, T. 2014. My artistic process: Haecceity series.

Hsiao, K.-A.; and Chen, L.-L. 2006. Fundamental dimensions of affective responses to product shapes. *International Journal of Industrial Ergonomics*, 36(6): 553–564. Kansei words.

Kang, Y.; Sun, Z.; Wang, S.; Huang, Z.; Wu, Z.; and Ma, X. 2021. MetaMap: Supporting Visual Metaphor Ideation through Multi-dimensional Example-based Exploration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, 1–15. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-8096-6.

Kim, B.; Wattenberg, M.; Gilmer, J.; Cai, C.; Wexler, J.; Viegas, F.; and Sayres, R. 2018. Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV). ArXiv:1711.11279 [stat].

Kobayashi, M.; Kinumura, T.; and Higashi, M. 2016. A Method for Supporting Aesthetic Design Based on the Analysis of the Relationships between Customer Kansei and Aesthetic Element. *Computer-Aided Design and Applications*, 13(3): 281–288. Kansei map, CAD design of chairs.

Langley, P. 2022. The Computational Gauntlet of Human-Like Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11): 12268–12273. *****, 2022.

Lévy, P.; and Yamanaka, T. 2009. Kansei Studies Description and Mapping through Kansei Study Keywords. *KANSEI Engineering International*, 8(2): 205–211.

Research, G.; and Projects, N. 2022. Mood Board Search by Google Research and Nord Projects.

Schmitt, P. 2019. The chair project (Four classics).

Schmitt, P.; and Weiß, S. 2018. The chair project: A case-study for using generative machine learning as automatism. In *ML for creativity and design workshop*. Montréal, Canada.

Sevener, Z. 2003. A semantic differential study of the influence of aesthetic properties on product pleasure. In *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, DPPI '03, 150–151. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-58113-652-4. Kansei words.

van der Stappen, A.; and Funk, M. 2021. Towards Guidelines for Designing Human-in-the-Loop Machine Training Interfaces. In *26th International Conference on Intelligent User Interfaces*, IUI '21, 514–519. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-8017-1. **.

Zhao, N.; Cao, Y.; and Lau, R. W. H. 2018. What characterizes personalities of graphic designs? *ACM Transactions on Graphics*, 37(4): 116:1–116:15. Kansei words.