Human Interaction and Emerging Technologies (IHIET 2023), Vol. 111, 2023, 377–387
https://doi.org/10.54941/ahfe1004044

The Effect of an AI Model for Conceptual Similarity on Design Ideation in a Co-Creative Design System

Jingoog Kim[1] and Mary Lou Maher[2]

[1]Loyola University New Orleans, New Orleans, LA 70118, USA
[2]University of North Carolina at Charlotte, Charlotte, NC 28223, USA

ABSTRACT

This paper describes a co-creative system that enhances design creativity in the initial idea generation process. The Collaborative Ideation Partner (CIP) is a co-creative design system that selects and presents inspirational images based on their conceptual similarity to the design task while the designer is sketching. In this paper, we present a study of how the different types of similarity of the contribution of the AI partner influences design ideation in a co-creative system. We conducted an experiment with a control condition in which the images are selected randomly from a curated database for inspiration and a treatment condition in which conceptual similarity is the basis for selecting the next inspiring image. Our results show that the AI model of conceptual similarity used in the treatment condition has a significant effect on the novelty, variety, and quantity of ideas during human design ideation.

Keywords: Human-computer co-creativity, Computational co-creative system, Ideation, Collaboration, Evaluation metrics

INTRODUCTION

Computational co-creative systems research is an emerging area that combines concepts from creativity support and autonomous creative systems in the broader field of computational creativity. Co-creative systems research has enormous potential since the concept can be applied to a variety of domains associated with creativity and encourage creative thinking. Understanding the effect of co-creative systems in the ideation process provides a basis for measuring the impact of AI on ideation across specific applications. The research in this paper shows how a co-creative agent influences the ideation process in a human-AI collaboration.

We present a co-creative design AI partner, the Collaborative Ideation Partner (CIP), that provides inspirational images based on their conceptual similarity to the design task. The AI model of CIP computes the conceptual similarity between the design task and the inspiring image using a curated image dataset and a pre-trained word2vec model. The turn-taking interaction between the user and the AI partner is designed to facilitate communication for design ideation. The CIP was developed to support an experiment that evaluates the effect of an AI model for conceptual similarity on design ideation in a co-creative design system.

In our previous studies, we explored the impact of AI-based inspiration based on visual and conceptual similarity and the temporal patterns on ideation using AI-based conceptual similarity inspiration (Kim et al., 2021a, 2021b; Kim and Maher, 2023). In this paper, we emphasize the effect of AI-based inspirations based on conceptual similarity. The contributions of this paper are (1) an AI model of conceptual similarity for determining the AI inspirations, and (2) the impact of inspirational images based on conceptual similarity on design ideation.

RELATED WORK

Co-creative systems in which users collaborate with an AI agent to make creative artifacts have been developed for creative domains such as art, music, dance, drawing, and games (Lin et al., 2020). Drawing Apprentice (Davis et al., 2015) is a web-based co-creative drawing system that analyzes the user's sketch and responds to the user's sketch. In the system, the user starts drawing a sketch on the canvas then the AI agent generates a sketch based on the users' sketch. DuetDraw (Oh et al., 2018) is an AI interface that allows users and the AI agent to draw pictures collaboratively such as completing the rest of the object that the user was drawing and automatically colorizing the sketches. Cobbie (Lin et al., 2020) is a mobile robot embedded with recurrent neural network (RNN)-based co-creative mobile drawing system that provides inspirational sketches under the command of the designer. Reframer (Lawton et al., 2023) is a co-creative drawing system in which the AI agent senses the user's drawing and uses CLIP-guided synthesis by optimization to determine the contributions of the AI. Reframer generates vector drawings from text prompts while other co-creative systems described above generate images based on the users' sketch. Creative Sketching Partner (Karimi et al., 2020) is a co-creative drawing system that provides inspirational sketches based on visual and conceptual similarity to the user's sketches. Our co-creative system is closely related to the Creative Sketching Partner, but our research specifically focuses on the impact of conceptual similarity on users' ideation, while the Creative Sketching Partner explores how a co-creative sketching system can guide users toward different types of creative design.

While previous research on co-creative systems claim that co-creative systems positively impact creative activities, assessing the impact of co-creative systems is still an ongoing research topic (Karimi et al., 2018). The evaluations of co-creative systems tend to focus on the interactive experience and the final product through evaluating the usability (Kim and Maher, 2021). Such studies use qualitative approaches (Kim and Maher, 2021) or a quantitative approach relying on questionnaires (Lawton et al., 2023; Lin et al., 2020) such as the Creativity Support Index (CSI) (Cherry and Latulipe, 2014). Our research focuses on how inspirations that an AI model provides based on conceptual similarity influence the designer's ideation process by measuring the novelty, variety, and quantity of ideas generated during design tasks.

COLLABORATIVE IDEATION PARTNER (CIP)

The Collaborative Ideation Partner (CIP) is a co-creative sketching system developed to explore the effect of an AI Model for conceptual similarity on ideation during a design sketching task. The user interface of the CIP is shown in Figure 1. There are two main spaces in the CIP interface: the design space (pink area) and the inspiring image space (purple area). The design space consists of a design task statement, undo button, clear button, and the user's canvas. The design task statement includes the name of the object to be designed as well as a context to further specify the objects' use and environment. The user can sketch on the user's canvas and edit the sketch using the undo and clear button. The inspiring image space includes an inspiring object name, "inspire me" button, and the AI partner's canvas. When the user clicks the "inspire me" button, the AI partner places an inspiring image in the AI partner's canvas. Ideation using CIP is a cyclical process in which the user sketches, asks for inspiration, and the AI partner presents an inspiring image. The title bar (grey area) of the CIP user interface includes a hamburger menu, the name of the system, and an introductory statement about the CIP system. The hamburger menu allows the selection from one of two design tasks, sink or bed, which allows the experiment facilitator to select one of the design tasks.

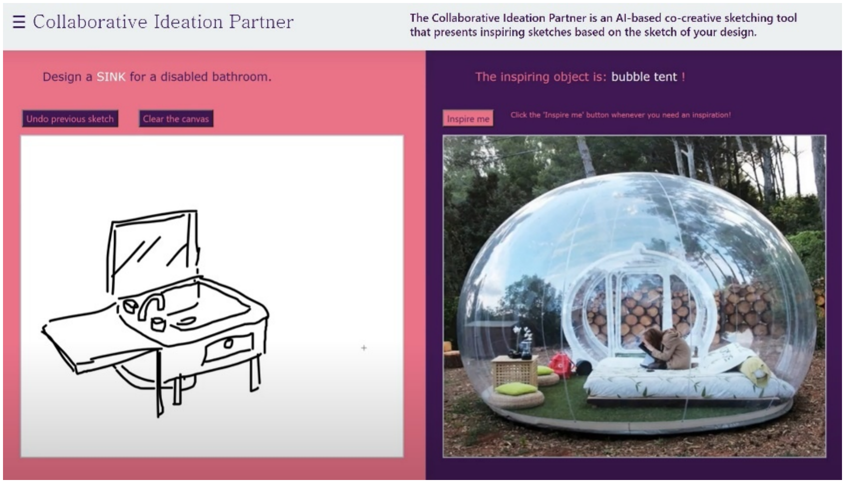


Figure 1: User interface of the collaborative ideation partner (Kim and Maher, 2023). https://www.amazon.com/Structure-Inflatable-Transparent-Rainproof-Windproof/dp/B0C6FBT8K1

DATASET

For the source of inspiring designs, we collected a dataset of high fidelity images of creative designs. To create the new dataset, we identified 20 object categories from the 345 categories sketches in the QuickDraw! dataset (Jongejan et al., 2016) based on their conceptual similarity to the object in the design task (sink and bed). We then searched for images of 5 creative designs online for each object category using keywords that combine "creative", "novel", "unusual", and "design" with the category of the design object (e.g. creative sink, unusual bed). The dataset thus contains 20 categories of objects with a total of 100 labeled images. Each image has three fields: id, object name, and design features. Id is the unique identifier that is assigned to each image. Object name is the name of the object that is represented in the image (e.g. electric massage bed, robotic advisor, smart sofa). Design features are keywords that represent the design features and unique functionalities of the design (e.g. multi-functional, entertainment, massage, combinational, digital, tv). The design features were assigned by the research team.

AI Model for Conceptual Similarity

The AI model for conceptual similarity uses a deep learning word embedding model to compute the degree of similarity between a set of words in the design task statement and a set of words for each image in the image dataset. We generate a pair-wise similarity score for each word in set 1 (words in the design task statement) and each word in set 2 (words in the design feature list for each image). A pre-trained word2vec model, trained on Wikipedia articles, is used to generate a vector representation for each of the words in both sets. We calculate the cosine similarity score for each pair of words for each image in the dataset. The similarity score for each image is calculated as the average of the pairwise cosine similarity scores. For example, a design task includes 4 words (i.e. bed, senior, living, facility) and an image includes 4 words of design features (e.g. comfort, massage, combinational, chair). For measuring the conceptual similarity between the design task and the image, we calculate the cosine similarity score for 16 pairs of words (4 words x 4 words) and then calculate the average of these 16 scores. We construct the conceptual similarity ranking for each image based on its similarity score. When the participant requests inspiration, the system uses the ranking in order from most conceptually similar to least conceptually similar to select the next image.

EXPERIMENT: MEASURING THE EFFECT OF CIP

The experiment for measuring the effect of CIP is designed to validate the following hypothesis:

- *AI-based conceptual similarity as the basis for inspiration increases the novelty, variety, and quantity of ideas during design ideation when compared to inspiration based on a random selection of relevant images.*

Study Design

The experiment is a within-subject design that compares participants' ideation while engaged in a design task with different ideation stimuli: a control condition with random inspirations (condition A), and a treatment condition with conceptually similar inspirations. All inspirations are selected from the curated dataset of 100 images of creative designs.

- Condition A (control condition): randomly selected inspiration (sink)
- Condition B (treatment condition): conceptually similar inspiration (bed)

During the study, for each participant and for each condition we collected video protocol data during the design session and a retrospective protocol after the design session. The protocol including the informed consent document has been reviewed and approved by our IRB and we obtained informed consent from all participants to conduct the experiment. We recruited 55 university students (N = 55) for the participants: each participant engaged in both conditions: a control condition (condition A) and a treatment condition (condition B).

The task is an open-end design task in which participants were asked to design an object in a given context through sketching. To reduce the learning effect, different objects for the design task were used for each condition: a sink for an accessible bathroom (condition A), and a bed for a senior living facility (condition B). The participants used a laptop and interacted with the CIP interface using a mouse to draw a sketch while performing the design task.

The procedure consists of a training session, two design task sessions, and two retrospective protocol sessions. In the training session, the participants are given an introduction to the features of the CIP interface and how to request inspiration from the AI partner. After the training session, the participants perform the two design tasks. The study used a counterbalanced order for the two design tasks. The participants have no time limits to complete the design task and were instructed to perform the design task until they were satisfied with their design. The participants are free to click the "inspire me" button as many times as they would like to get inspiration from the system. However, the participants were told to request at least 3 inspirational sketches, i.e. clicking the "inspire me" button at least 3 times during a single design task. Once the participants finish the two design task sessions, the participants are asked to explain what they were thinking based on watching their design session recording as time goes on, and how the AI's sketches inspired their design in the retrospective protocol session.

Data Segmentation and Coding

Two types of data were collected for analyzing the experiment's results: a set of sketches that participants produced during the design tasks and verbalizing the ideation process during the retrospective protocol. We recorded the design task sessions and retrospective sessions for each participant. The sketch data shows the progress of the design ideation and the final design visually for each design task session. The verbal data describes how the participants came up with ideas collaborating with the AI partner and applied the ideas to their design.

To analyze the verbal data collected from the retrospective sessions, the verbal data of all retrospective protocol sessions was transcribed. The transcripts were segmented based on the inspiring images the participant clicked. An inspiring image segment starts when the participant requests an inspiring image and ends when the participant requests the next inspiration. In this study, we define an idea as a cognitive issue using the FBS ontology (Gero and Kannengiesser, 2004). We further segmented the inspiring segments until each segment has a single code in the FBS ontology. An inspiring segment thus includes multiple idea segments. After segmenting the verbal data, we conducted stemming, the process of reducing inflected words to their word stem. This stemming process allows us to identify unique ideas and repeat ideas in a design session. Two coders coded the 110 sessions of retrospective protocol (i.e. 55 sessions of condition A, 55 sessions of condition B) individually based on FBS ontology and then came to a consensus for the different coding results.

ANALYSIS OF CODED DATA

To measure ideation in the design sessions, we developed three metrics based on the measurement of ideation in (Shah et al., 2003): novelty, variety, and quantity of design. These metrics provide a basis for evaluating the effect of AI inspiration on exploring (variety and quantity) and expanding (novelty) the design space (Shah et al., 2003).

The transcription data comprises 110 sessions of retrospective protocol. In condition A, the participants had a total of 704 inspiring images and produced 4,226 ideas. In condition B, the participants had a total of 583 inspiring images and produced 5,673 ideas. The participants had more inspiring images in condition A but produced more ideas in condition B, as shown in Table 1. This result indicates that AI inspiration based on conceptual similarity is more effective in design ideation than random inspiration.

Table 1. Total number of inspirations participants clicked and total number of ideas participants generated.

#Inspiration		#Ideas generated	
Condition A	Condition B	Condition A	Condition B
704	583	4,226	5,673

Novelty

Novelty is a measure of how unusual or unexpected an idea is as compared to other ideas (Shah et al., 2003). In our experiment, a novel idea is defined as a unique idea across all design sessions in one design task. For measuring novelty, we count how many novel ideas are in the entire collection of ideas in a design session then divide the novel ideas by the number of inspirations that the designer gets from a co-creative system, as shown in Equation (1). The novelty score thus means the number of novel ideas per AI inspiration in a design session.

$$Novelty\ Score\ = \frac{\Sigma(Unique\ ideas\ across\ all\ design\ sessions\ in\ a\ condition)}{\Sigma(Inspirations\ in\ a\ design\ session)} \tag{1}$$

The participants produced 356 novel ideas in condition A and 480 novel ideas in condition B. The result of the novelty score showed that 36 participants revealed a higher novelty score in condition B than in condition A

(Figure 2). A paired t-test was conducted to determine the significance of the result between the control condition and the treatment condition in novelty. The results showed a significant difference between the control condition and the treatment condition. Participants in condition B ($M = 1.04$, $SD = 0.96$) produced higher novelty scores than in condition A ($M = 0.82$, $SD = 0.88$), $t(54)=-2.28$, two tail $p = 0.029016$. This result indicates that AI inspiration is more effective in producing more novel ideas than when co-creating ideas with random inspirations.

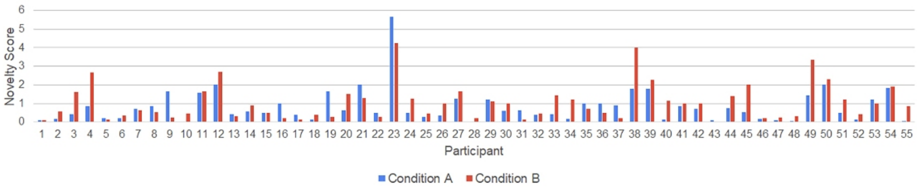


Figure 2: Novelty scores of participants.

In addition to the analysis of novelty scores, we observed the video stream data of retrospective protocol sessions to see how the participants generate novel ideas communicating with the AI-based inspirations. From the observation of the video stream data, we discovered that participants in Condition B tended to initiate idea development from inspiring images and generate novel ideas by expanding upon the concept of inspiring images. On the other hand, participants in Condition A were more likely to produce ideas that were unrelated to the inspiring images. Figure 3 shows an example of a novel idea and how the participant responds to the inspiring image to generate a novel idea. The inspiring image in Figure 3 is a robotic operating room. From this image, many participants noticed the robotic arm, display, and hanging lights as a common idea. However, P50 produced a novel idea by transferring the robotic arm concept to a robotic bed which has robotic legs for mobility. P50 then started to expand the function of the walking robot by attaching additional features. P50 added a robotic arm with a big umbrella and a light that can support staying in an outdoor environment such as shading and lighting.

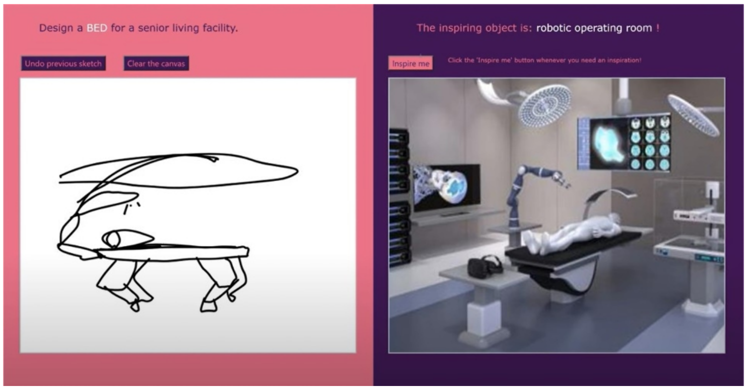


Figure 3: Example of novel idea that P50 generated. https://www.engadget.com/2016-11-28-hospital-to-get-first-dedicated-3d-tissue-printing-facility.html

Variety

Variety is a measure of the explored solution space during the idea generation process (Shah et al., 2003). For measuring variety, we code each idea as a new idea or a repeated idea in a design session and only the number of new ideas is counted in a design session. For example, if an idea of "side table" appeared four times in a design session, we coded the idea as a new idea for the first appearance and coded the other three appearances as a repeated idea. After coding new and repeated ideas, we divide the new ideas by the number of inspirations that the designer gets from a co-creative system, as shown in Equation (2).

$$\textit{Variety Score} = \frac{\Sigma(\textit{New ideas in a design session})}{\Sigma(\textit{Inspirations in a design session})} \tag{2}$$

The participants produced 2,046 new ideas in condition A and 2,944 new ideas in condition B. The result of the variety score showed that 40 participants revealed a higher variety score in condition B than in condition A, as shown in Figure 4. A paired t-test was conducted to determine the significance of the result between the control condition and the treatment condition in variety. The results showed a significant difference in variety. Participants in condition B ($M = 6.60$, $SD = 4.51$) produced higher variety scores than in condition A ($M = 4.62$, $SD = 3.83$), $t(54)=-4.42$, two tail $p = 0.000046$. This result indicates that AI inspiration is more effective in producing a variety of ideas than when co-creating ideas with random inspirations.

From the observation of the video stream data, we found that participants tend to generate various ideas associated with the details of the design and the context of the design getting inspiration from the CIP. Figure 5 shows the final designs P38 produced in condition A and condition B, along with the variety score for each condition: condition A: 10.80, and condition B: 18.67. While P38 in condition A mostly focused on the sink object itself with simple representations of the major features, P38 in condition B showed considered more design details (i.e. side table, leg part of the bed, seating part of the bed, head part of the bed) and the context (i.e. book, plant, coffee cup, vase).

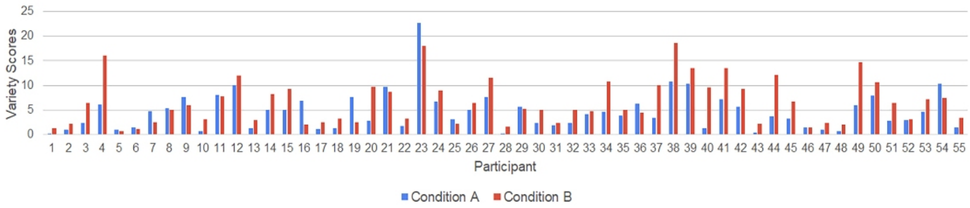
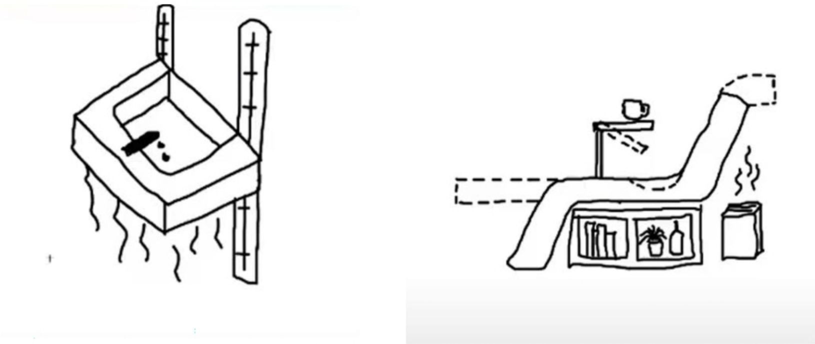


Figure 4: Variety scores of participants.



(a) Condition A (variety score: 10.80) (b) Condition B (variety score: 18.67)

Figure 5: Example of variety of ideas that P38 generated.

Quantity

Quantity is the total number of ideas generated. According to Shah et al. (Shah et al., 2003), generating more ideas increases the possibility of better ideas. For measuring quantity, the number of ideas, including both new ideas and repeated ideas is counted in a design session, as shown in Equation (3).

$$\text{Quantity Score} = \frac{\Sigma(\text{New ideas in a design session}) + \Sigma(\text{Repeated ideas in a design session})}{\Sigma(\text{Inspirations in a design session})} \tag{3}$$

The participants produced 4,226 ideas in condition A and 5,673 ideas in condition B. The result of the quantity score showed that 41 participants revealed a higher quantity score in condition B than in condition A, as shown in Figure 6. The result of a paired t-test showed a significant difference in quantity. Participants in condition B ($M = 12.43$, $SD = 8.54$) produced higher quantity scores than in condition A ($M = 9.48$, $SD = 8.59$), $t(54) = -3.10$, two tail $p = 0.002994$. This result indicates that AI inspiration is more effective in increasing the quantity of ideas than when co-creating ideas with random inspirations.

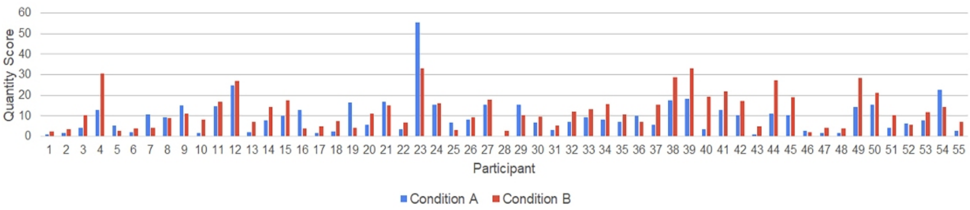


Figure 6: Quantity scores of participants.

In examining the result of quantity scores, some participants showed a large difference between the 2 conditions: P55 had a quantity score in condition A: 2.77 and in condition B: 7.04. Figure 7, showing the final designs P55 produced in condition A and condition B, highlights how P55 designed an entire room (i.e. accessible bathroom, bed room in a senior living facility) rather than designing the design object (sink or bed). A common design pattern of P55 in both conditions is to include many individual items such as chair, table, remote control, and plant. However, the design in condition B shows a much larger number of items than the design in condition A.

(a) Condition A (quantity score: 2.77) (b) Condition B (quantity score: 7.04)

Figure 7: Example of quantity of ideas that P55 generated.

CONCLUSION

This paper describes a co-creative design system called Collaborative Ideation Partner (CIP) that provides inspirational images during a design task based on the conceptual similarity to the task description. To study the impact of CIP on ideation, we performed an experiment with two conditions for the inspiration: random inspiration and conceptually similar inspiration. To evaluate the effect of AI inspiration, we measured the ideas generated by the user with three metrics: novelty, variety, and quantity. We developed an approach for measuring ideation that is based on a segmentation of verbal protocol data according to the cognitive concepts of FBS and quantitative metrics for novelty, variety, and quantity of ideas. Our results show that AI-based conceptual similarity as the basis for inspiration increases the novelty, variety, and quantity of ideas during design ideation. In addition to the statistical analysis of the individual novelty, variety, and quantity scores, we found that the participants tend to request more inspiring images in the control condition but produce more ideas in the treatment condition. Our method for measuring ideation can be used more generally for comparing the impact of different AI models on the designer's ideation process.

ACKNOWLEDGMENT

We acknowledge Safat Siddiqui for his assistance in coding the protocol data.

REFERENCES

Cherry, E., Latulipe, C., 2014. Quantifying the creativity support of digital tools through the creativity support index. ACM Transactions on Computer-Human Interaction (TOCHI) 21, 1–25.

Davis, N., Hsiao, C.-Pi., Singh, K. Y., Li, L., Moningi, S., Magerko, B., 2015. Drawing apprentice: An enactive co-creative agent for artistic collaboration, in: Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition. pp. 185–186.

Gero, J. S., Kannengiesser, U., 2004. The situated function–behaviour–structure framework. Design studies 25, 373–391.

Jongejan, J., Rowley, H., Kawashima, T., Kim, J., Fox-Gieg, N., 2016. The quick, draw!-ai experiment. Mount View, CA, accessed Feb 17, 4.

Karimi, P., Grace, K., Maher, M. L., Davis, N., 2018. Evaluating creativity in computational co-creative systems. arXiv preprint arXiv:1807.09886.

Karimi, P., Rezwana, J., Siddiqui, S., Maher, M. L., Dehbozorgi, N., 2020. Creative sketching partner: an analysis of human-AI co-creativity, in: Proceedings of the 25th International Conference on Intelligent User Interfaces. pp. 221–230.

Kim, J., Maher, M. L., 2023. The effect of AI-based inspiration on human design ideation. International Journal of Design Creativity and Innovation 1–18.

Kim, J., Maher, M. L., 2021. Evaluating the Effect of Co-Creative Systems on Design Ideation., in: ICCC. pp. 440–443.

Kim, J., Maher, M. L., Siddiqui, S., 2021a. Studying the Impact of AI-based Inspiration on Human Ideation in a Co-Creative Design System., in: IUI Workshops.

Kim, J., Maher, M. L., Siddiqui, S., 2021b. Collaborative Ideation Partner: Design Ideation in Human-AI Co-creativity., in: CHIRA. pp. 123–130.

Lawton, T., Ibarrola, F. J., Ventura, D., Grace, K., 2023. Drawing with Reframer: Emergence and Control in Co-Creative AI, in: Proceedings of the 28th International Conference on Intelligent User Interfaces. pp. 264–277.

Lin, Y., Guo, J., Chen, Y., Yao, C., Ying, F., 2020. It is your turn: collaborative ideation with a co-creative robot through sketch, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14.

Oh, C., Song, J., Choi, J., Kim, S., Lee, S., Suh, B., 2018. I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–13.

Shah, J. J., Smith, S. M., Vargas-Hernandez, N., 2003. Metrics for measuring ideation effectiveness. Design studies 24, 111–134.