Studying the Impact of AI-based Inspiration on Human Ideation in a Co-Creative Design System

Jingoog Kim[a], Mary Lou Maher[a] and Safat Siddiqui[a]

[a] *University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA*

Abstract

Co-creative systems in design enable users to collaborate with an AI agent on open-ended creative tasks in the design process. This paper describes a co-creative system that supports design creativity by encouraging the exploration of design solutions in the initial idea generation process. The Collaborative Ideation Partner (CIP) is a co-creative design system that provides inspirational sketches based on the visual and conceptual similarity to sketches drawn by a designer. To evaluate the effect of CIP on design ideation, we conducted an exploratory study that measures ideation in a co-creative system. To measure the ideation, we developed a way of measuring ideation in a co-creative system including an outcome and a process approach. From the exploratory study, we learned that the image quality in the dataset is important in AI-based creativity and inspirations based on conceptual similarity to the target design have more impact on ideation than inspirations based on visual similarity to sketches drawn by a designer. We present the architecture of the CIP system and a study design based on what we learned from the exploratory study.

Keywords [1]

Co-Creativity, Co-Creative System, Ideation, Collaboration

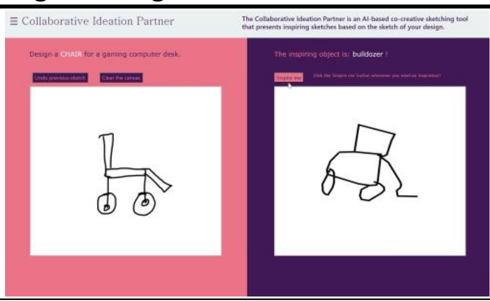
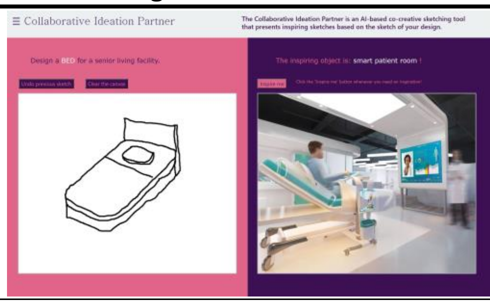
1. Introduction

Computational co-creative systems are a growing research area in computational creativity. While some research on computational creativity has a focus on generative creativity [1]–[9], co-creative systems focus on how systems that implement generative creativity can work with humans on a creative task [10]–[17]. Co-creative systems have enormous potential to enhance human creativity since they can be applied to a variety of domains associated with creativity and encourage designers' creative thinking. Understanding the effect of co-creative systems in the ideation process can aid in the design of the generative AI models in co-creative systems and the evaluation of the impact of co-creative systems on human creativity.

We present a co-creative sketching AI partner, the Collaborative Ideation Partner (CIP), that provides inspirational sketches based on the visual and conceptual similarity to sketches drawn by a designer. To select an inspiring sketch, the AI model of CIP computes the visual similarity of images in a data set based on the vector representations of visual features of the sketches and the conceptual similarity based on the category names of the sketches using two pre-trained word2vec models. 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 exploratory study that evaluates the effect of an AI model for visual and conceptual similarity on design ideation in a co-creative design tool.

[1]Joint Proceedings of the ACM IUI 2021 Workshops, April 13-17, 2021, College Station, USA
EMAIL: jkim168@uncc.edu (A. 1); M.Maher@uncc.edu (A. 2); ssiddiq6@uncc.edu (A. 3)
ORCID: 0000-0003-3597-6153 (A. 1)

Table 1
Comparison of original design and current design of CIP

	Original Design of CIP	Current Design of CIP
Interface		
Stimuli	Visual and Conceptual Similarity	Conceptual Similarity
Inspiring images	Low fidelity sketches of a general object	high fidelity images of a creative design
Modes of inspiration	4 modes: random, similar, conceptually similar and visually different, visually similar and conceptually different	2 modes: random inspiration, conceptually similar
Dataset	3450 Sketches (QuickDraw [18])	100 images

To evaluate the impact of co-creative systems in design, we measure design ideation in a co-creative system. Ideation, an idea generation process for conceptualizing a design solution, is a key step that can lead a designer to an innovative design solution in the design process. Idea generation is a process that allows designers to explore many different areas of the design solution space [19]–[24]. Ideation has been studied in human design tasks and collaborative tasks in which all participants are human. Collaborative ideation can help people generate more creative ideas by exposing them to ideas different from their own [25]. This paper has a focus on evaluating how a co-creative agent influences the ideation process in a human-AI collaboration.

In this paper, we describe an exploratory study measuring ideation when co-creating with the CIP and what we learned from the exploratory study. To measure ideation in a co-creative system, we employ two approaches: an outcome-based approach that focuses on the end product of the design, and a process-based approach that focuses on thought processes during the design. From the exploratory study, we learned that the quality of the images in the dataset is important in AI-based creativity for the impact on designer's creativity and inspirations based on conceptual similarity to the target design has more impact on ideation than inspirations based on visual similarity to sketches drawn by a designer. We updated the CIP system and study design based on what we learned from the exploratory study. Table 1 shows the comparison between the CIP system used for the exploratory study and the current CIP system based on what we learned from the exploratory study. The current CIP system focuses on conceptually similar inspirations to the target design and provides high fidelity images of creative designs. This study aims to identify the effect of AI inspiration on design ideation through a way of measuring ideation in a co-creative system.

2. Computational co-creative systems

Computational co-creative systems are one of the growing fields in computational creativity that involves human users collaborating with an AI agent to make creative artifacts. The distinction of co-creativity from computational creativity is that co-creativity is a collaboration in which multiple parties contribute to the creative process in a blended manner [26]. Co-creative systems have been applied in different creative domains such as art, music [14], dance [15,27], drawing [28], and game design [39,56].

Evaluating co-creative systems is still an open research question and there is no standard metric for measuring computational co-creativity [31].

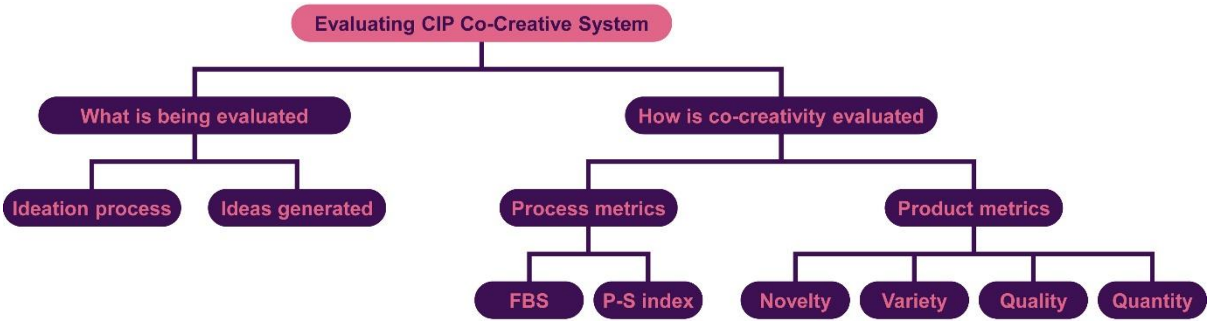


Figure 1: Evaluating the creative ideation partner

The research on co-creative systems shows various approaches to evaluate co-creative systems and computational co-creativity. Some researches focus on evaluating the interactive experience [29,14,30,31] and others focus on the effectiveness of the system to produce or generate a creative outcome [39,56]. Karimi et al. [31] presented a framework for evaluating creativity in computational co-creative systems. This framework responds to four questions that serve to characterize the many and varied approaches to evaluating computational co-creativity: who is evaluating the creativity, what is being evaluated, when does evaluation occur, and how the evaluation is performed. The framework enables comparisons of evaluation focus and methods across existing co-creative systems. Using this framework, we have shown that the evaluations of the existing co-creative systems described in this section respond to "what is being evaluated" with a focus on evaluating the interactive experience and the final product. In this paper, we respond to "what is being evaluated" and "how is the evaluation performed" by evaluating the ideation process using FBS and Problem-Solution index, and the metrics for evaluating the novelty, variety, quality, and quantity of ideas in the creative outcome, as shown in Figure 1. Section 3 describes how we define and measure ideation in more detail.

3. Defining and measuring design ideation

Ideation is a creative process where designers generate, develop, and communicate new ideas. Ideation in design can lead to innovative design solutions through generating diverse concepts [19]–[22], [24]. The goal of design is to develop useful and innovative solutions and design ideation allows designers to explore different areas of the design solution space [23,32]. A design process is an evolution of different kinds of representations [33]. In a design process, designers externalize and visualize their design intentions and communicate with visualizations to interact with their internal mental images [34]. During ideation, designers commonly use freehand sketches and rough physical models as a tool for constructing external representations as cognitive artifacts of design [35]. Making sketches and physical models is an interaction, a conversation [36]. In the ideation stage, designers frame problems producing new discoveries through the conversation. The graphical and physical representations as cognitive artifacts are essential in the ideation process.

Many ideation methods have been developed to support designers in generating innovative design solutions. Ideation methods provide a normative procedure on how to overcome certain blocks to creativity [37]. Analogy is an ideation method and we focus on analogy to develop a co-creative design tool. Analogical reasoning is an inference method in design cognition to develop a design leading to unexpected discoveries [38]. Design-by-Analogy (DbA) is a design tool that provides inspiration for innovative design solutions. Inspirations in Design-by-Analogy (DbA) are achieved by transferring a design problem (source) to a solution (target) in another domain [39]. The association between a source design and a target design can be based on semantic (conceptual) characteristics or visual (structural) representations. The semantic and visual stimuli thus can be a basis for developing computational tools that support design ideation. The Collaborative Ideation Partner (CIP), a co-creative design system we present in this paper, uses visual and conceptual similarity metrics as key factors for collaborative ideation using design by analogy.

Evaluation of ideation can be classified into two groups: outcome-based approaches and process-based approaches [40]. Outcome-based approaches focus on evaluating the ideation process based on the designs, or outcomes, and the characteristics of ideas generated. Process-based approaches focus on evaluating idea generation processes based on the cognitive processes inherent to creative thought. Process-based approaches collect data via a protocol study and analysis using ideation cognitive models. Outcome-based approaches have become more prevalent than process-based approaches due to the inherent complexity and difficulties in using process-based approaches [41]. There have been several metrics used to evaluate the performance of idea generation techniques such as fluency and novelty that cognitive psychologists consider as the primary measures of idea generation. Shah et al. [41] introduced four types of outcome-based metrics for measuring ideation effectiveness that are commonly used for evaluating idea generation in design: novelty, variety, quality, and quantity of designs. Novelty is a measure of how unusual or unexpected an idea is as compared to other ideas. Variety is a measure of the explored solution space during the idea generation process. The generation of similar ideas indicates low variety and hence, less probability of finding better ideas in other areas of the solution space. Quality is a subjective measure of the feasibility of an idea and how close it comes to meet the design specifications. Quantity is the total number of ideas generated, generating more ideas increases the possibility of better ideas. These metrics enable a comparison of a designer's exploration of a design space using different ideation methods.

Process-based approaches evaluate idea generation based on the cognitive processes via a protocol analysis and cognitive models. The Function-Behavior-Structure (FBS) ontology [42,43] is a design ontology that describes designed things, or artifacts, irrespective of the specific discipline of designing. The function (F) of a designed object is defined as its teleology; the behavior (B) of that object is either derived (Bs) or expected (Be) from the structure, where structure (S) represents the components of an object and their compositional relationships. These ontological classes are augmented by requirements (R) that come from outside the designer and description (D) that is the document of any aspect of designing. In this ontological view, the goal of designing is to transform a set of requirements and functions into a set of design descriptions. The transformation of one design issue into another is defined as a design process [44].

The design process can be viewed as interactions between two notional design spaces: problem space and solution space [45,46]. The Problem-Solution (P-S) index [47,48] is a measurement capturing the meta-level structures of design cognition in terms of problem-focused and solution-focused design issues. This measurement uses an integration of the FBS ontologically-based coding scheme with a Problem-Solution (P-S) division [47,48]. In the P-S division, design issues of R, F, and Be are mapped to problem space and design issues of Bs, and S are mapped to solution space [47]. A design session with a P-S index larger than 1 as one with a problem-focused designing style, and a session with a P-S index value less than or equal to 1 as one with a solution-focused style. The P-S index can be used to compare design cognition while using different creativity techniques for concept generation in collaborative design settings.

4. The collaborative ideation partner (CIP)

We developed the Collaborative Ideation Partner (CIP) as a co-creative design system which builds on previous works [49,50] that interprets sketches drawn by a user and provides inspirational sketches based on visual similarity and conceptual similarity. We developed the CIP to explore evidence for the hypothesis that: AI models for contributions to a creative product based on a measure of visual and conceptual similarity produce different ideation processes and outcomes than the random condition.

The user interface of CIP is shown in Figure 2. There are two main spaces in the CIP interface: the drawing space (pink area) and the inspiring sketch space (purple area). The drawing space consists of a design task statement, undo button, clear button, and user's canvas.

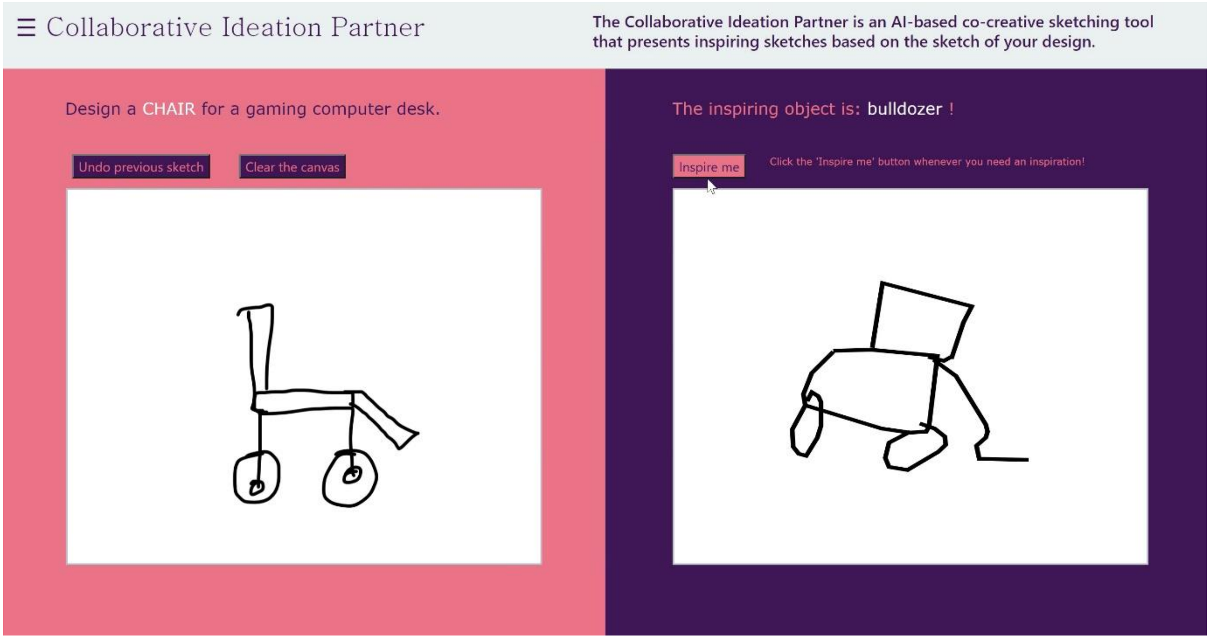


Figure 2: User interface of collaborative ideation partner

The design task statement in the drawing space includes the object to be designed as well as a context to further specify the objects' use and environment. The user can draw a sketch in the drawing space and edit the sketch using the undo and clear button. The inspiring sketch space includes an "inspire me" button, the name of the inspiring object, and a space for presenting the AI partner's sketch. When the user clicks the "inspire me" button after sketching their design concept, the AI partner provides an inspiring sketch based on visual and conceptual similarity. An ideation process using CIP involves turn-taking communications between the user and the AI partner. Another part of the CIP interface in addition to the two main spaces is the top area (grey area) including a hamburger menu and an introductory statement. The hamburger menu on the top-left corner of the interface includes four design tasks (i.e. sink, bed, table, chair) and allows the experiment facilitator to select one of the design tasks. Each design task is associated with different categories of ideation stimuli.

4.1. Dataset

For the source of inspiring sketches, the original CIP uses a public benchmark dataset called QuickDraw! [18], which was created during an online game where players were asked to draw a particular object within 20 seconds. The dataset includes 345 categories with more than 50 million labeled sketches, where sketches are the array of the x and y coordinates of the strokes. The system uses the simplified drawing json files that use Ramer–Douglas–Peucker algorithm [51,52] to simplify the strokes, and position and scale the sketches into a 256 X 256 region. The stroke data associated with these sketches are used to calculate the visual similarity and the corresponding category names are used to measure the conceptual similarity.

4.2. AI models for visual and conceptual similarity

The CIP has 2 distinct components for measuring similarity between the user's sketch and the sketches in the dataset: one component for calculating visual similarity and another component for calculating conceptual similarity. Figure 3 shows how the CIP system identifies an inspiring sketch: the visual similarity is based on the vector representations of visual features of the sketches and the conceptual similarity based on the category names of the sketches using two pre-trained word2vec models.

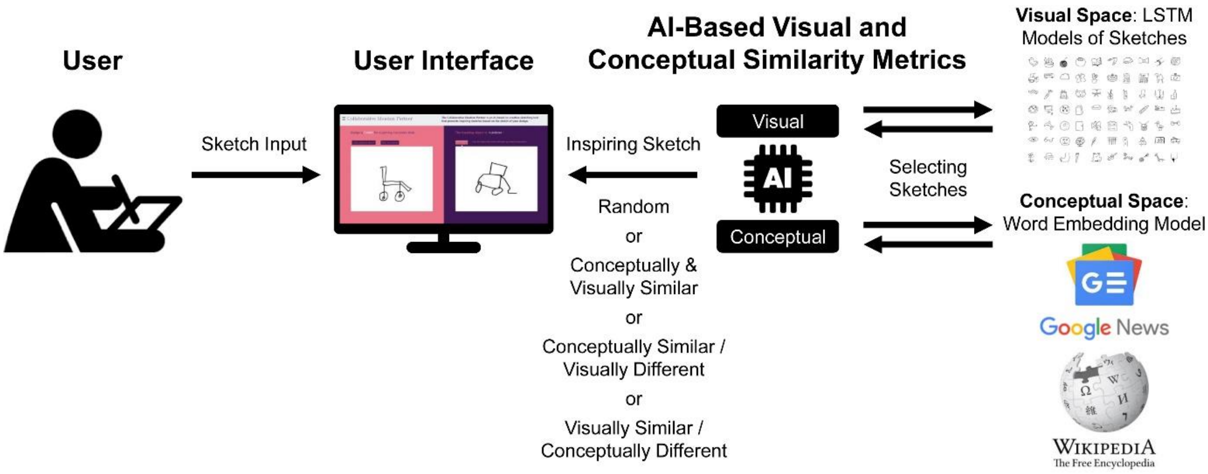


Figure 3: AI-Based Co-Creativity in the CIP

For the visual similarity component, we followed the precedent for using neural network models in computational creativity described in [50,53] and trained a model with 3 convolutional layers, 2 LSTM layers, and a softmax output layer on the QuickDraw dataset. This model provides a latent space representation for measuring the distance, or similarity, between 2 sketches. For all the sketches in the dataset, we collected the last LSTM layer of the trained model and used that as the vector representations of visual features of the sketches. We used the K-means algorithm to identify 10 clusters of sketches and randomly selected one sketch of each cluster as a typical sketch for that cluster of sketches. Thus, we converted the QuickDraw dataset of 50 million sketches into 3450 sketches (345 categories, each has 10 sketches). To prepare the user's sketch for comparison with the sketches in the dataset, the CIP collects the user sketches as an array of x and y coordinates of strokes and simplifies the strokes using Ramer–Douglas–Peucker algorithm [51,52]. It also positions and scales the user's sketch into the 256 X 256 region to match the sketch format with the input dataset of the trained model. The CIP takes the last LSTM layer of the trained model as the vector representation of visual features of the user's sketch, and calculates the Euclidean distance to measure visual similarity between the user's sketch and 3450 sketches of Quickdraw dataset. The visual similarity component of the CIP prepares a sorted list of visually similar sketches to generate the final sequence of sketches in the conceptual component of CIP that considers the conceptual similarities of the sketches.

For the conceptual similarity component, we considered sketch category names in the QuickDraw dataset as the concepts of the sketches that contain 345 unique categories. We used two pre-trained word2vec models, Google News [54] and Wikipedia [55], and calculated cosine similarities for measuring the conceptual similarities between the object categories of the design tasks and the categories of inspiring sketches from the dataset. For each category of the design tasks, we generated two sorted lists of conceptually similar category names, one for each word2vec model, and then used human judgement to compare the sorted lists and select the top 15 common conceptually similar category names that appear in both lists. This final step of using human judgement improved the alignment between the conceptual similarities of AI models and human perception. The conceptual similarity component of CIP uses the common list of category names for sorting the sketches based on the conceptual similarities.

4.3. AI-based inspiration in CIP

To support an exploratory study that measures ideation when co-creating with CIP, the interaction with CIP has four distinct modes of inspiration that vary the visual and conceptual similarity. Each of the four modes appears as a design task (i.e. sink, bed, table, chair) in the CIP interface.

- Inspire with a random sketch (sink): The CIP selects a sketch randomly from the sketch dataset to be displayed on the AI partner's canvas.
- Inspire with a visually and conceptually similar sketch (bed): The CIP selects a sketch from a set of sketches where each one is similar visually and conceptually to the user's sketch (e.g. user sketch - a bed, AI sketch - a similar shape of bed to the user's sketch).
- Inspire with a conceptually similar and visually different sketch (table): The CIP selects a sketch from a set of sketches where each one is conceptually similar but visually different to the user's sketch (e.g. user sketch - a square table, AI sketch - a round table).
- Inspire with a visually similar and conceptually different sketch (chair): The CIP selects a sketch from a set of sketches where each one is visually similar but conceptually different to the user's sketch (e.g. user sketch - a circular chair back, AI sketch - a face).

5. Exploratory study: measuring ideation when co-creating with the CIP

The goal of the exploratory study is to evaluate the effect of AI inspiration on ideation through an analysis of the correlation between conceptual and visual similarity with characteristics of ideation. To measure ideation when co-creating with the collaborative ideation partner, we applied both evaluation methods of ideation: an outcome-based approach (i.e. novelty, variety, quality, quantity) and a process-based approach (i.e. P-S index).

5.1. Study design

The type of study is a mixed design of between-subject and within-subject design with four conditions. There are 3 groups of within-subject design (i.e. A&B, A&C, A&D) in this study and each group has a control condition (i.e. condition A) and one of 3 treatment conditions (i.e. condition B, C, D).

- Condition A (control condition): randomly (sink)
- Condition B (treatment condition): visually and conceptually similar (bed)
- Condition C (treatment condition): conceptually similar and visually different (table)
- Condition D (treatment condition): visually similar and conceptually different (chair)

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 12 students from human-centered design courses for the participants: each participant engaged in 2 conditions: a control condition and one of the treatment conditions, with 4 participants for each of the 3 groups of within-subject design (i.e. A&B, A&C, A&D). The experiment is a mixed design with N=4 and a total of 12 participants.

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 a accessible bathroom (condition A), a bed for a senior living facility (condition B), a table for a tinkering studio, a collaborative space for designing, making, building, crafting, etc. (condition C), a chair for a gaming computer desk (condition D). 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 they work to enable the AI partner to provide inspiration during their design task. After the training session, the participants perform two design tasks in a control condition and a treatment condition. The study used a counterbalanced order for the two design tasks. The participants have no time limits to complete the design task. The participants were given as much time as needed 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 have at least 3 inspirational sketches (i.e. clicking the "inspire me" button at least 3 times during a design session), a minimum number of inspirations, from the system. The facilitator is present during the design task but does not interfere in the design process. 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.

5.2. Observations of ideation with CIP

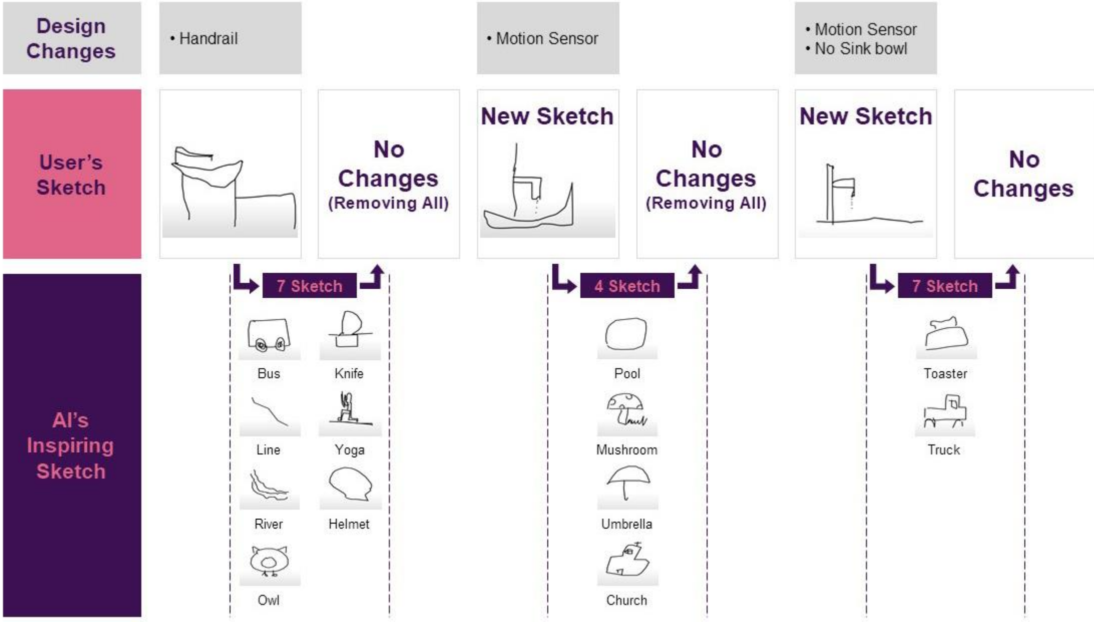
We observed the video stream data to see how participants develop their design ideas communicating with the inspirations and the participants' responses to inspirations show different patterns of users on the use of CIP in an ideation process. Figure 4 shows typical examples of the process for the evolution of the participant's sketch using CIP in each condition. In an evolution of the participant's sketch, participants in each condition start with a basic shape of the target design then develop the design with inspiration from the AI partner. Participants explored many inspiring sketches in condition A but did not have many design changes; while participants in conditions B, C, and D developed their design in response to fewer inspiring sketches.

As shown in Figure 4a, P1 drew a basic sink with a handrail before getting the first inspiration then tried to get an inspiration from the AI partner. P1 had 7 inspiring sketches but did not change anything for the design. P1 then cleaned all the canvas then drew a new sketch which is a sink with a motion sensor. P1 had 4 inspiring sketches and did not change anything again for the design. P1 cleaned the canvas and drew a new sketch again applying the motion sensor idea again then had 2 inspiring sketches. However, P1 finally finished the design without any changes. During the retrospective session, P1 mentioned he did not get ideas from the inspiring sketches several times, for example *"I don't have any inspiration with the pictures."* This case shows an example that participants do not have many ideas from random inspirations.

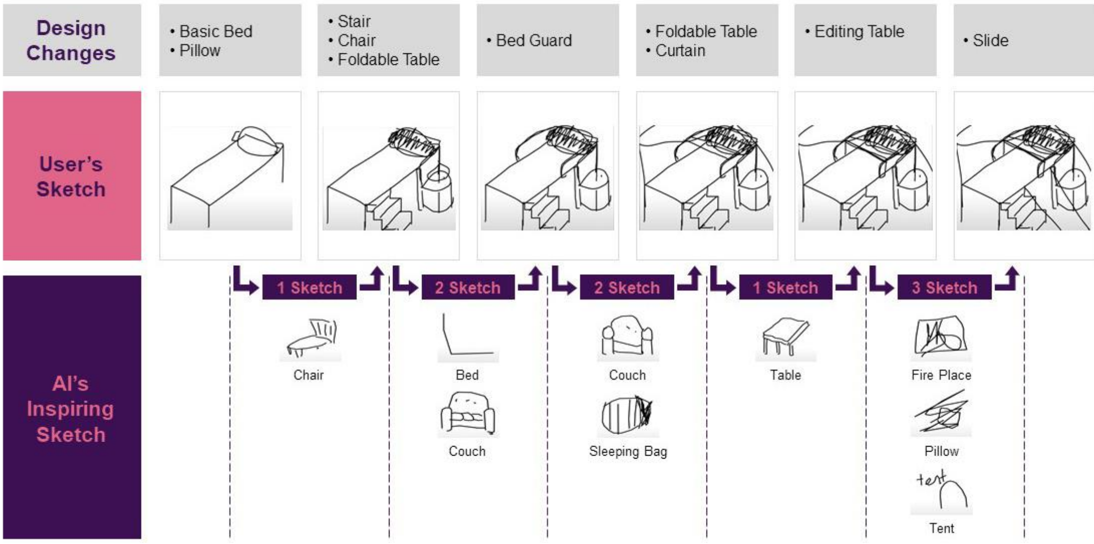
As shown in Figure 4b, P4 drew a basic bed and a pillow before getting the first inspiration then requested inspiration from the AI partner. The first inspiring sketch was a chair and P4 added a stool, table, and a stair next to the bed. After that, P4 had two more inspirations, bed and couch, and added bed guard around the bed. P4 described that the bed guard idea came from the armrest of the couch. P4 then had 2 more inspiring sketches, couch and sleeping bag, and added a curtain. P4 mentioned that the curtain idea came from the enclosing feature of the sleeping bag and couch. The next inspiring sketch is a table and P4 edited the foldable table on the bed. After that, P4 had three more inspiring sketches and added a slide that helps getting out of the bed easily. P4 described that the slide idea came from the shape of the tent.

As shown in Figure 4c, P2 drew a rectangle for a table before getting the first inspiration then tried to get an inspiration from the AI partner. The first inspiring sketch was a golf club and P2 added table legs mimicking the shape of a golf club. P2 then had a fireplace sketch and added a large grid paper on the table. P2 described that the grid paper idea came from the way the lines are drawn in the fireplace. After that, P2 had matches and added a table lamp. P2 then had two pool sketches and added a pencil cup. The last inspiring sketch is a wine glass but P2 did not change the design with the wine glass.

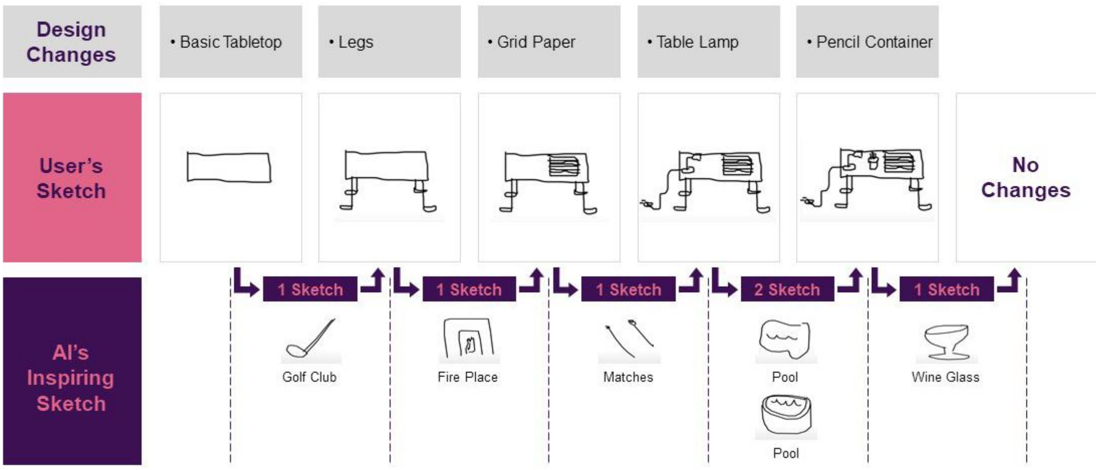
As shown in Figure 4d, P3 drew a basic chair without any special function for the context of gaming before getting the first inspiration then requested inspiration from the AI partner. The first inspiring sketch was a raccoon and P3 added an ear shape decoration on the top of the chair and an eye shape headrest getting an inspiration from the shape of the raccoon sketch (i.e. ear, and eye). P3 described that *"So, I saw the raccoon and I kind of liked how its ears were. Cause I have seen things, where people have really interesting chairs, and I think people that game may usually want more interesting chairs. So, I thought it'd be cool to have little ears at the top, and then make the mask kind of like, a pillow."* After that, P3 had the second inspiring sketch which is a power outlet.



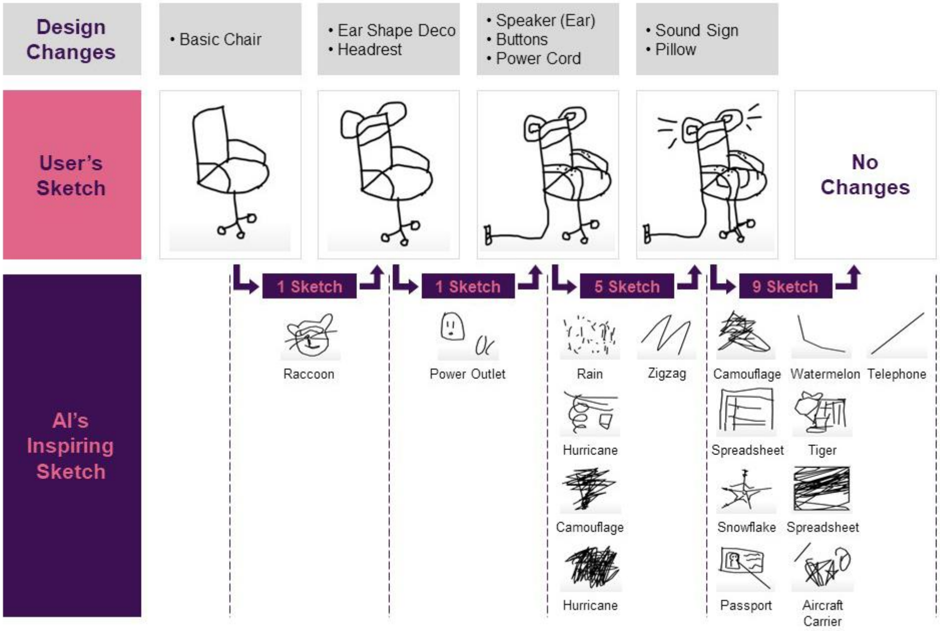
(a) The evolution of P1 design in condition A (randomly)



(b) The evolution of P4 design in condition B (visually and conceptually similar)



(c) The evolution of P2 design in condition C (conceptually similar and visually different)



(d) The evolution of P3 design in condition D (visually similar and conceptually different)

Figure 4: AI-Based Co-Creativity in the CIP

P3 added a speaker on the ear decoration, buttons on the armrest to control sound volume/massage/lights, and power cord. P3 described that *"the power outlet really gave me a lot of the inspiration. I thought... instead of just random ears, it could be like a speaker. Then, I thought all the little dots on the armrest could be buttons, to do things. If it's different things, like sound volume, or could do a massage. The little line coming out of it, would be to plug it into the wall, so all the buttons could work."* In this case, the idea came from the inspiring sketch was transferred to new functions of the chair while the idea came from the raccoon was transferred to the shape of the chair. P3 then had 5 more inspiring sketches (i.e. rain, hurricane, zigzag, and camouflage). P3 mentioned that they were inspired from the irregular lines of the sketches and added sound projecting lines next to the ear shape speaker and a pillow on the seat. After that, P3 had nine more inspiring sketches, but did not change anything for the chair design. P3 talked about a decoration idea from the star shape of an aircraft carrier and snowflake but did not change the chair design.

5.3. Data collected

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

5.4. Data segmentation and coding

To analyze the verbal data collected from the retrospective sessions, the verbal data of all retrospective protocol sessions (i.e. 12 sessions of condition A, 4 sessions of condition B, 4 sessions of condition C, and 4 sessions of condition D) was transcribed. The transcripts were segmented based on the inspiring sketches the participant clicked. A segment starts with an inspiring sketch and ends when the inspiration is clicked for the next sketch. To identify each idea in an inspiring sketch segment, we segmented the inspiring segments again based on FBS ontology [42,43] as an idea segment, since an inspiring sketch segment includes multiple ideas. An inspiring segment thus includes multiple idea segments. The idea segments were coded based on FBS ontology [42,43] as requirement (R), function (F), expected behavior (Be), behavior from structure (Bs), and structure (S). A R segment is an utterance that talks about the given requirement in the statement of design task (e.g. accessible bathroom) or a new requirement the participant came up with for the design (e.g. if someone is not able to reach the height); a F segment is an utterance that talks about a purpose or a function of the design object (e.g. more accessible); a Be segment is an utterance that talks about an expected behaviors from the structure (e.g. water could automatically come out), a Bs segment is an utterance that talks about a behavior derived from the structure (e.g. pressing on), a S segment is an utterance that talks about a component of the design object (e.g. button). Two coders coded the idea segments individually based on the coding scheme above then came to consensus for the different coding results.

5.5. Measuring ideation: outcome-based approach

For the outcome-based approach, we developed four metrics based on [41]: novelty, variety, quality, and quantity of design. Novelty is a measure of how unusual or unexpected an idea is as compared to other ideas. In this study, a novel idea is defined as a unique idea across all design sessions in a condition. For measuring novelty, we counted how many novel ideas in the entire collection of ideas in a design session (personal level of novelty) and a condition (condition level of novelty). We removed the same ideas across all design sessions in a condition then counted the number of ideas. Variety is a measure of the explored solution space during the idea generation process. Each idea segment was coded whether it is a new idea or a repeated idea in a design session. For measuring variety in this study, only the number of new ideas coded as R/F/B/S is counted in a design session while the metric of quantity includes both new ideas and repeated ideas. Quality is a subjective measure of the design. In this study, quality is measured using the Consensual Assessment Technique (CAT) [56], a method in which a panel of expert judges is asked to rate the creativity of projects. Two judges, researchers involved in this study, individually evaluated the final design in each condition as low/medium/high quality, in two evaluation rounds. In the first-round of evaluation, each judge evaluated the final designs identifying some criteria for evaluating the quality of ideas. Once the judges finished the first-round of evaluation, they shared the criteria they identified/used, not sharing the results of the evaluation, then made a consensus for the criteria that will be used for the second-round evaluation. The criteria that the judges agreed for evaluating the quality of ideas in this study are the number of features, how responsive the features are to the specific task, how creative the design is. In the second-round evaluation, each judge evaluated the final design again using the agreed criteria. Quantity is the total number of ideas generated. For measuring quantity in this study, the number of ideas both new ideas and repeated ideas coded as R/F/B/S is counted in a design.

5.6. Measuring ideation: process-based approach

For the process-based approach, we used the P-S index [48] to examine the design cognition from a meta-level view (i.e., a single-value measurement). For the meta-level view, the P-S index is calculated by computing the number of the total occurrences of the design issues concerned with the problem space (i.e. R, F, Be) and related to the solution space (i.e. Bs, S). A design session with a P-S index larger than 1 as one with a problem-focused style, and a session with a P-S index value less than or equal to 1 as one with a solution-focused style. In addition to calculating the P-S index of each design session, we looked at the number of problems and solutions to identify a distinct difference between the conditions.

5.7. What we learned from exploratory study

With the data collected in the exploratory study, we compared outcome-based features (i.e. novelty, variety, quality, quantity) and process-based features (i.e. P-S index). Our findings show that the AI-based stimuli produce different ideation outcomes and processes when compared to random stimuli.

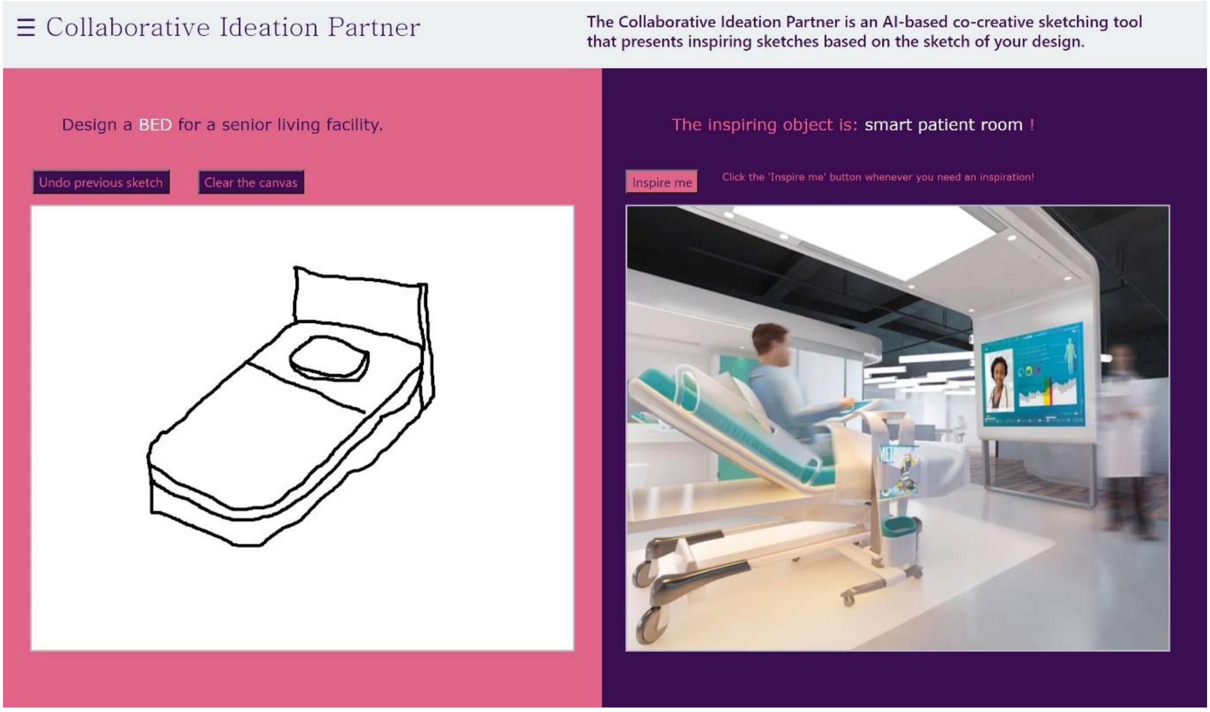


Figure 5: Current CIP user interface

Novel ideation, evidenced by an increase in the variety and quantity of ideas, is associated with AI-based conceptually similar stimuli. The findings from analysis of the P-S index show that AI-based visual and conceptual similarity is associated with a problem-focused designing style that produces more solutions than we found in the condition with random inspirations. We found that participants in condition C (conceptually similar and visually different stimuli) produced more functions than in condition A (random).

In our observations of the exploratory study, we identified some issues on the sketch data set and AI-based visually similar stimuli. First, the quality of the sketch dataset is very important to inspire participants to come up with new ideas. The sketches in this dataset are not the result of a design process. The sketches in the QuickDraw dataset are generated to represent the basic shape of a given well known object. Based on the retrospective protocol data, participant's ideas mostly came up from purposes, functions, features, and structures of the inspiring sketches, and the simple representation of objects in the QuickDraw dataset were not providing very rich inspiration. Second, the complexity of participants' sketches increased during the design session, affecting the accuracy of the visual similarity measure used to select an inspiring sketch. The AI model for visually similarity to the participant's sketch was more accurate at the beginning of the design session, but was less accurate as the participant's sketch became more complicated. Third, the CIP in condition D (visually similar and conceptually different) often provides sketches that are not visually similar to the participant's sketch since the inspiring sketches are first selected to be conceptually different, and that reduces the potential for identifying sketches that are visually similar.

6. Current CIP and study design

From the exploratory study we learned that inspiration based on conceptual similarity has more impact on the novelty and variety of ideas than visual similarity and that the quality of the dataset is important for the design ideation. For our current CIP and study design, we developed a more comprehensive model for conceptual similarity based on multiple features of the design rather than only a categorical word, and collected a dataset of designs as the basis for inspiration rather than a dataset of sketches on well known objects. Figure 5 shows the current CIP user interface providing an inspiring image of a creative design instead of a simple sketch. The target design is a bed for a senior living facility and the inspiring image is a smart patient room. The smart patient room is the most conceptually similar design to the target design. The design of the smart patient room includes many functions and objects associated with the context of a senior living facility such as reclining bed, bed table, magazine holder, trash can, digital screen for health care, and wheels for mobility.

6.1. Dataset

For the source of inspiring designs, we collected a dataset of high fidelity images of creative designs. To create the new dataset, we selected 20 common categories from the categories of QuickDraw dataset that are conceptually similar to the target designs of the exploratory study (i.e. sink, bed, table, chair). We then searched for images of 5 creative designs online for each category using keywords “creative”, “novel”, “unusual”, “design” (e.g. creative sink, unusual bed). The dataset thus contains 20 categories with 100 labeled images. Each image has three fields: id, object name, and design feature. Id is the unique identifier that is assigned to each image. Object name is the name of the design that is represented in the image (e.g. electric massage bed, robotic advisor, smart sofa). Design feature is keywords that represent the design features and unique functionalities of the design (e.g. multi-functional, entertainment, massage, combinational, digital, tv).

6.2. AI model for conceptual similarity

The AI model for conceptual similarity computes the degree of similarity between a set of words in the design task and a set of words for each image in the image dataset. While the previous CIP system used the object category of the design task (e.g. bed) to measure the conceptual similarity, the updated CIP used a set of words in the design task statement (i.e. bed, senior, living, facility) to include the context of the design object for measuring the conceptual similarity. For measuring conceptual similarity, we thus use the words in the design task statement (i.e. bed, senior, living, facility) and the words in the design features of each image in the image dataset. We generate a pair-wise similarity score for each word in set 1 (i.e. words in the design task statement) and each word in set 2 (i.e. words in the design feature). A Wikipedia pre-trained word2vec model is used to calculate the similarity between the two words using a pair-wise comparison: a word from the design task statement and a word from the design features of an image. We calculate the cosine similarity score for each pair of a design task statement word and a design feature word and create a set of cosine similarity scores including all pairs of design task statement words and design task feature words for each image in the image dataset. As a conceptual similarity score between the target design and the image, we use the average score of cosine similarity scores for each image. For example, a design statement 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 target design and the image, we calculate each cosine similarity score for 16 pairs of words (4 words x 4 words) then calculate the average cosine similarity. We create the conceptual similarity ranking based on the similarity score of each image. The system selects from the most conceptually similar image in order when the user clicks the inspire button.

6.3. Study design

In our study design we focus on the impact of the AI model for conceptual similarity on design ideation. The experimental conditions include a control condition and one treatment condition. The experiment focuses on identifying distinct patterns of the participant's ideation in a human-AI collaboration where the AI partner contributes content based on conceptual similarity. The experiment is a within-subject design that compares participants’ ideation outcome and process while engaged in a design task with different ideation stimuli: a control condition with random inspirations (condition A), a treatment condition with conceptually similar inspirations.

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

We recruited 50 university students (N=50) for the participants: each participant engaged in

2 conditions: a control condition (condition A) and a treatment condition (condition B). We use two design tasks (i.e. condition A: design a sink for an accessible bathroom, condition B: design a bed for a senior living facility) that was used for the exploratory study. The data collection includes the video of the design sessions and video of the retrospective protocol sessions, as in the exploratory study. The data from this study is still being collected.

7. Conclusion

This paper presents a co-creative design tool called Collaborative Ideation Partner (CIP) that supports the idea generation of new designs with stimuli that vary in similarity to the user's design in two dimensions: conceptual and visual similarity. The AI models for measuring similarity in the CIP use deep learning models and cosine similarity to the user's sketch and design task. The interactive experience allows the user to seek inspiration as needed. To study the impact of varying levels of visual and conceptual similar stimuli, we performed an exploratory study with four conditions for the AI inspiration: random, high visual and conceptual similarity, high conceptual similarity with low visual similarity, and high visual similarity with low conceptual similarity.

We developed an approach for measuring ideation that has two components: an outcome-based approach and a process-based approach. The outcome-based approach adapts existing quantitative metrics for ideation: novelty, variety, quality, and quantity of ideas expressed in the outcome. The process-based approach uses existing cognitive models of design, including the FBS ontology and the P-S index, to code and analyze the verbal protocol of the designers. These measures can be used in evaluating the impact of AI contributions in other co-creative systems that support design creativity. We applied these measures in the four conditions in the CIP to demonstrate how to operationalize our approach for measuring ideation in a co-creative system.

From the exploratory study, we learned that the quality of dataset is important in AI-based creativity for the impact on designer's creativity and inspirations based on conceptual similarity to the target design leads to more novel ideation than inspirations based on visual similarity to sketches drawn by a designer. We updated the CIP system and study design based on what we learned from the exploratory study. The current CIP system focuses on conceptually similar inspirations to the target design and provides high fidelity images of creative designs.

8. References

[1] Y.-G. Cheong and R. M. Young, Narrative generation for suspense: Modeling and evaluation, in Joint International Conference on Interactive Digital Storytelling, 2008, pp. 144–155.

[2] S. Colton, J. Goodwin, and T. Veale, Full-FACE Poetry Generation., in ICCC, 2012, pp. 95–102.

[3] M. Cook and S. Colton, Ludus Ex Machina: Building A 3D Game Designer That Competes Alongside Humans., in ICCC, 2014, pp. 54–62.

[4] L. A. Gatys, A. S. Ecker, and M. Bethge, A neural algorithm of artistic style, arXiv preprint arXiv:1508.06576, 2015.

[5] F. Rashel and R. Manurung, Pemuisi: a constraint satisfaction-based generator of topical Indonesian poetry., in ICCC, 2014, pp. 82–90.

[6] T. Veale, Coming good and breaking bad: Generating transformative character arcs for use in compelling stories, 2014.

[7] T. Veale and Y. Hao, A fluid knowledge representation for understanding and generating creative metaphors, in Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008), 2008, pp. 945–952.

[8] G. A. Wiggins, Searching for computational creativity, New Generation Computing, vol. 24, no. 3, pp. 209–222, 2006.

[9] G. A. Wiggins, A preliminary framework for description, analysis and comparison of creative systems, Knowledge-Based Systems, vol. 19, no. 7, pp. 449–458, 2006.

[10] J. A. Biles, Interactive GenJam: Integrating real-time performance with a genetic algorithm, 1998.

[11] K. Compton and M. Mateas, Casual Creators., in ICCC, 2015, pp. 228–235.

[12] N. Davis, E. Y.-L. Do, P. Gupta, and S. Gupta, Computing harmony with PerLogicArt: perceptual logic inspired

collaborative art, in Proceedings of the 8th ACM conference on Creativity and cognition, 2011, pp. 185–194.

[13] N. M. Davis, Human-computer co-creativity: Blending human and computational creativity, 2013.

[14] G. Hoffman and G. Weinberg, Gesture-based human-robot jazz improvisation, in 2010 IEEE International Conference on Robotics and Automation, 2010, pp. 582–587.

[15] M. Jacob, A. Zook, and B. Magerko, Viewpoints AI: Procedurally Representing and Reasoning about Gestures., 2013.

[16] T. Lubart, How can computers be partners in the creative process: classification and commentary on the special issue, International Journal of Human-Computer Studies, vol. 63, no. 4–5, pp. 365–369, 2005.

[17] B. Magerko, C. Fiesler, A. Baumer, and D. Fuller, Bottoms up: improvisational micro-agents, in Proceedings of the Intelligent Narrative Technologies III Workshop, 2010, pp. 1–8.

[18] J. Jongejan, H. Rowley, T. Kawashima, J. Kim, and N. Fox-Gieg, The quick, draw!-ai experiment, Mount View, CA, accessed Feb, vol. 17, p. 2018, 2016.

[19] Ö. Akin, Necessary conditions for design expertise and creativity, Design Studies, vol. 11, no. 2, pp. 107–113, 1990.

[20] C. J. Atman, J. R. Chimka, K. M. Bursic, and H. L. Nachtmann, A comparison of freshman and senior engineering design processes, Design studies, vol. 20, no. 2, pp. 131–152, 1999.

[21] D. R. Brophy, Comparing the attributes, activities, and performance of divergent, convergent, and combination thinkers, Creativity research journal, vol. 13, no. 3–4, pp. 439–455, 2001.

[22] N. Cross, Design cognition: Results from protocol and other empirical studies of design activity, in Design knowing and learning: Cognition in design education, Elsevier, 2001, pp. 79–103.

[23] S. R. Daly, S. Yilmaz, J. L. Christian, C. M. Seifert, and R. Gonzalez, Design heuristics in engineering concept generation, 2012.

[24] Y.-C. Liu, A. Chakrabarti, and T. Bligh, Towards an ‘ideal’approach for concept generation, Design studies, vol. 24, no. 4, pp. 341–355, 2003.

[25] J. Chan et al., Semantically far inspirations considered harmful? accounting for cognitive states in collaborative ideation, in Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition, 2017, pp. 93–105.

[26] L. Mamykina, L. Candy, and E. Edmonds, Collaborative creativity, Communications of the ACM, vol. 45, no. 10, pp. 96–99, 2002.

[27] M. Jacob, G. Coisne, A. Gupta, I. Sysoev, G. G. Verma, and B. Magerko, Viewpoints ai, 2013.

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

[29] P. Lucas and C. Martinho, Stay Awhile and Listen to 3Buddy, a Co-creative Level Design Support Tool., in ICCC, 2017, pp. 205–212.

[30] G. N. Yannakakis, A. Liapis, and C. Alexopoulos, Mixed-initiative co-creativity, 2014.

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

[32] A. Newell and H. A. Simon, Human problem solving, vol. 104. Prentice-Hall Englewood Cliffs, NJ, 1972.

[33] V. Goel and P. Pirolli, The structure of design problem spaces, Cognitive science, vol. 16, no. 3, pp. 395–429, 1992.

[34] T. Dorta, Design flow and ideation, International Journal of Architectural Computing, vol. 6, no. 3, pp. 299–316, 2008.

[35] W. Visser, The cognitive artifacts of designing. CRC Press, 2006.

[36] T. Dorta, E. Perez, and A. Lesage, The ideation gap:: hybrid tools, design flow and practice, Design studies, vol. 29, no. 2, pp. 121–141, 2008.

[37] N. V. Hernandez, J. J. Shah, and S. M. Smith, Understanding design ideation mechanisms through multilevel aligned empirical studies, Design Studies, vol. 31, no. 4, pp. 382–410, 2010.

[38] J. S. Gero and M. L. Maher, Mutation and analogy to support creativity in computer-aided design, 1991.

[39] D. P. Moreno et al., Fundamental studies in Design-by-Analogy: A focus on domain-knowledge experts and applications to transactional design problems, Design Studies, vol. 35, no. 3, pp. 232–272, 2014.

[40] B. A. Nelson, J. O. Wilson, D. Rosen, and J. Yen, Refined metrics for measuring ideation effectiveness, Design Studies, vol. 30, no. 6, pp. 737–743, 2009.

[41] J. J. Shah, S. M. Smith, and N. Vargas-Hernandez, Metrics for measuring ideation effectiveness, Design studies, vol. 24, no. 2, pp. 111–134, 2003.

[42] J. S. Gero, Design prototypes: a knowledge representation schema for design, AI magazine, vol. 11, no. 4, pp. 26–26, 1990.

[43] J. S. Gero and U. Kannengiesser, The situated function–behaviour–structure framework, Design studies, vol. 25, no. 4, pp. 373–391, 2004.

[44] J. S. Gero, Generalizing design cognition research, DTRS, vol. 8, pp. 187–198, 2010.

[45] K. Dorst and N. Cross, Creativity in the design process: co-evolution of problem–solution, Design studies, vol. 22, no. 5, pp. 425–437, 2001.

[46] M. Maher and H.-H. Tang, Co-evolution as a computational and cognitive model of design, Research in Engineering design, vol. 14, no. 1, pp. 47–64, 2003.

[47] J. S. Gero, H. Jiang, and C. B. Williams, Design cognition differences when using unstructured, partially structured, and structured concept generation creativity techniques, International Journal of Design Creativity and Innovation, vol. 1, no. 4, pp. 196–214, 2013.

[48] H. Jiang, J. S. Gero, and C. Yen, Exploring designing styles using a problem–solution index, 2014.

[49] N. Davis, S. Siddiqui, P. Karimi, M. L. Maher, and K. Grace, Creative Sketching Partner: A Co-Creative Sketching Tool to Inspire Design Creativity., in ICCC, 2019, pp. 358–359.

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

[51] D. H. Douglas and T. K. Peucker, Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, Cartographica: the international journal for geographic information and geovisualization, vol. 10, no. 2, pp. 112–122, 1973.

[52] U. Ramer, An iterative procedure for the polygonal approximation of plane curves, Computer graphics and image processing, vol. 1, no. 3, pp. 244–256, 1972.

[53] P. Karimi, M. L. Maher, N. Davis, and K. Grace, Deep Learning in a Computational Model for Conceptual Shifts in a Co-Creative Design System, arXiv preprint arXiv:1906.10188, 2019.

[54] T. Mikolov, K. Chen, G. Corrado, and J. Dean, Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781, 2013.

[55] R. Rehurek and P. Sojka, Software framework for topic modelling with large corpora, 2010.

[56] T. M. Amabile, Social psychology of creativity: A consensual assessment technique., Journal of personality and social psychology, vol. 43, no. 5, p. 997, 1982.