THE COGNITIVE STUDY OF DESIGN IDEATION IN AN AI-BASED CO-CREATIVE SKETCHING PARTNER

by

Jingoog Kim

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Software and Information Systems

Charlotte

2022

Approved by:

____
Dr. Mary Lou Maher

____
Dr. John Gero

____
Dr. Doug Markant

____
Dr. Gabriel Terejanu

ABSTRACT

JINGOOG KIM. Cognitive study of design ideation in an AI-based co-creative sketching partner. (Under the direction of DR. MARY LOU MAHER)

The primary goal of design is to provide effective and innovative solutions for solving design problems. Ideation, an initial idea generation for conceptualizing a design solution, is a key step that can lead design 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. Due to the importance of ideation, many studies focused on understanding the cognitive processes in idea generation and evaluating ideation. This thesis focuses on the idea generation process based on conceptual similarity in a human-AI collaboration. Co-creative systems in design allow users to collaborate with an AI agent on open-ended creative tasks in the design process. Co-creative systems share the characteristics of both creativity support tools helping users achieve creative goals and algorithms that generate creative content autonomously. Co-creative systems support design creativity by encouraging the exploration of design solutions in the initial idea generation. However, there is a lack of studies about the effect of co-creative systems on the cognitive process during ideation. This thesis posits that the contribution of an AI partner in design is associated with specific properties of ideation such as novelty, variety, quality, and quantity of ideas.

This thesis presents 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. This thesis addresses how the conceptual similarity of the contribution of the AI partner influences design ideation in a co-creative system. This thesis presents 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. To evaluate the ideation during the use of CIP, this thesis employed an aggregate analysis and a temporal analysis. The findings 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.

DEDICATION

This dissertation work is dedicated to my dearest wife and my beloved daughter, Lina Lee and Erin Kim, who always love me and lead me through the valley of darkness with the light of hope and support. I am truly thankful for having you in my life. Love you forever. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my dissertation. I also dedicate this dissertation to all my family members in South Korea. A special feeling of gratitude to my loving mother, Soon Im Bae who encouraged me to pursue my dreams and complete my dissertation. Thanks for your love and support.

ACKNOWLEDGEMENTS

I am grateful to everyone who has helped me in my struggle to achieve my dream of becoming a Ph.D. I would first like to thank my advisor, Dr. Mary Lou Maher, for her consistent support and guidance throughout the writing of this dissertation. Your insightful feedback critically helped me to refine my thinking about designing and evaluating co-creative systems.

I would also like to thank the rest of my dissertation committee: Dr. John Gero, Dr. Doug Markant, and Dr. Gabriel Terejanu. Thanks for your invaluable advice, continuous support, and patience during my Ph.D. study that made this thesis so much better.

I would like to acknowledge my colleague, Safat Siddiqui, who helped develop the Collaborative Ideation Partner system and analyze the user study data. The work presented in this thesis would not have been possible without your effort. Finally, I would like to express my gratitude to all the members of the HCI Lab, who have provided friendship and support, and with whom I have shared laughter, frustration, and companionship. It is their kind help and support that have made my study and life in the United State a wonderful time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
CAT	Consensual Assessment Technique
DbA	Design-by-Analogy
FBS	Function-Behavior-Structure
HCI	Human Computer Interaction
IRB	Institutional Review Board
IRR	Inter-Rater Reliability
LSTM	Long Short-Term Memory
P-S	Problem-Solution
RNN	Recurrent Neural Network
UX	User Experience

CHAPTER 1: INTRODUCTION

Computational co-creative systems research is an emerging focus that combines concepts from creativity support and autonomous creative systems in the broader field of computational creativity. While some research on computational creativity focus on generative creativity [9, 13, 16, 31, 72, 81, 82, 84, 85], co-creative systems focus on computer systems collaborating with humans on a creative task [5, 15, 20, 24, 41, 43, 44, 55, 58]. 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. While co-creative systems can be applied to a variety of domains associated with creativity and encourage designers' creative thinking, there are few studies that focus on evaluating co-creative systems. Most research on co-creative systems focuses on evaluating the usability and the interactive experience [49] rather than how the co-creative system influences creativity in the creative process. To evaluate the usability and the user experience of co-creative systems, the studies often used qualitative approaches and a few studies have used a quantitative approach to evaluate the user experience of co-creative systems relying on questionnaires [48] such as the System Usability Scale (SUS) [6] and the Creativity Support Index (CSI) [10]. Understanding the effect of co-creative systems in the ideation process can aid in the design of co-creative systems and evaluation of the effectiveness of co-creative systems.

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 [1, 3, 7, 17, 54]. 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 [8]. This thesis begins to address how a co-creative agent influences the ideation process in a human-AI collaboration.

Sketching is a major tool to externalize and visualize design ideas or communicate with them during the ideation stage. This thesis presents 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. This thesis emphasizes the effect of the AI-based inspirations based on the conceptual similarity.

1.1 Research Motivation

AI abilities are becoming more and more competent. We're seeing humans interacting with AI, more than ever. AI is now becoming part of our lives. When people research human-AI interaction, they focus on the user experience and the final outcome. There's a lack of studies on how it affects the way we think. As a UX designer and an architect, I have used many design tools and saw that the abilities of design tools often impact designers' creativity and the final outcomes. Computational design tools such as creativity support tools allow designers to create complex designs we could not create without computers, and to create designs easily with the embedded design sources in design tools. The abilities of computational design tools thus changed the way designers traditionally think. This realization inspires me to study a computational co-creative partner that provides creative inspiration in human-AI collaboration. Now we are at another inflection point with AI. We need to understand how AI changes the way we think. This thesis is looking at an important topic, humans in AI are the future. It is a new paradigm of designing that is certainly going to have AI.

Given an understanding of the academic landscape of human-AI interaction and computational co-creativity, my dissertation work focused on designing, developing, and evaluating a co-creative design tool, the Creative Ideation Partner (CIP). Along the way, I studied computational co-creative systems and ideation to inform the design and evaluation of the Creative Ideation Partner (CIP) system. However, since computational co-creativity is a new research domain and there is no standard metric to evaluate computational co-creativity [49], existing metrics used in human ideation were adapted to evaluate co-creative systems. While the main contribution of this thesis is the design and evaluation of the Creative Ideation Partner (CIP), the evaluation method and cognitive studies presented in this thesis also have significant value to the study of co-creation more broadly.

1.2 Thesis Statement

Co-creative agents in design can support design creativity by providing inspirations based on conceptual similarity in the ideation process. During the idea generation, the inspirational designs based on the conceptual similarity can influence distinct properties in the cognitive process. This thesis explores the effect of an AI model for conceptual similarity through evaluating design ideation for co-creative systems.

Thesis statement: *In a co-creative system, the conceptual similarity of the contribution of the AI partner can enhance the novelty, variety, and quantity of the design ideas in design ideation. Using the aggregate analysis and the temporal analysis of ideation as a theoretical framework to quantify the contribution of an AI partner on human-AI co-creation can help objectively evaluate the effect of AI-based inspirations in co-creative systems.*

1.3 Research Questions and Hypotheses

This research asks the overarching research question: How do AI-based inspirations impact human ideation in a creative design task and how can we design more effective co-creative systems for designing? It explores the effect of AI-based inspirations on design ideation to understand how co-creative agents can influence the design ideation process in human-AI collaboration. This research addresses this thesis statement by contributing to the following research questions:

- Research Question 1: *Are AI-based inspirations based on conceptual similarity more effective than random inspirations in design ideation when measuring ideation effectiveness with the metrics of novelty, variety, and quantity?*
- Research Question 2: *What are the patterns of novelty, variety, quantity in human ideation when providing AI-based inspirations based on conceptual similarity?*

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

- Hypothesis 1: *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.*
- Hypothesis 2: *The quantity in novelty, variety, and quantity of ideas over time decreases more slowly in an ideation with AI-based inspirations based on conceptual similarity than the temporal pattern of ideation with random inspirations in a creative design task.*

1.4 Methodology

This thesis uses a mixed-method approach to address and explore the research questions and hypotheses outlined above. There are three main activities comprising the methods: 1) designing and developing a co-creative system that supports human ideation in open-ended creative design tasks, 2) developing an approach to measuring ideation as a basis for evaluating co-creative systems in design, and 3) applying the measuring approach to explore and identify the effect of co-creative systems in design.

The Collaborative Ideation Partner (CIP) is a co-creative sketching system which builds on previous research [22, 50] that interprets sketches drawn by a user and provides inspirational sketches based on visual similarity and conceptual similarity. There are two versions of the CIP: CIP Sketch which is an initial version of CIP and CIP Design which is an updated version of CIP. The CIP Sketch is a co-creative design system that provides inspirational sketches based on the visual and conceptual similarity to sketches drawn by a designer. The CIP Sketch was developed to support an exploratory study that explores the effect of an AI model for visual and conceptual similarity on design ideation in a co-creative design tool. Based on what we learned from the exploratory study, the design of CIP Sketch was changed to CIP Design. The CIP Design focuses on conceptually similar inspirations to the target design and provides high fidelity images of creative designs.

This thesis presents a method of measuring ideation as a basis for evaluating the effect of AI models on the design process. To develop a way of measuring ideation in a co-creative system, I studied several approaches related to evaluating co-creative systems and measuring ideation. The approach to evaluating co-creative systems and measuring ideation in co-creative systems includes defining the effect of the co-creative system in design, defining a design idea in design ideation using a co-creative system, metrics to measure the effectiveness of ideation, and approaches to evaluate the ideation in human-AI collaboration. In this thesis, the effect of the co-creative system is defined as contributions of the AI agent to the idea generation. We define an idea as a cognitive issue that the designer considers during the design process, and adopt the Function-Behavior-Structure (FBS) ontology [32, 35] as a basis for segmenting and coding each idea in the design process. In the exploratory study using the CIP Sketch, we developed the four metrics based on [76] to analyze the coded data of the retrospective protocol session: novelty, variety, quality, and quantity of design ideas. The exploratory study employed two approaches: an outcome-based approach and a process-based approach. In the user study of the CIP Design, we updated the metrics and the approach to measure ideation. The updated metrics include novelty, variety, and quantity of design ideas. The user study of the CIP Design focuses on a cognitive-based approach rather than a product-based approach and employs two approaches with the three metrics: an aggregated approach and a temporal approach.

I performed two studies (i.e. an exploratory study of the CIP Sketch and a user study of the CIP Design) to apply the evaluation method and to investigate the effect of the CIP system on design ideation. To evaluate the CIP Sketch, an exploratory study was conducted 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. The verbal data was collected from 24 retrospective sessions (i.e. 12 participants' verbal data, N=4) were analyzed. To evaluate the CIP Design, a user study was conducted 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. The verbal data was collected from 110 retrospective sessions (i.e. 55 participants' verbal data, N=55) were analyzed.

1.5 Contributions

The main contributions of this dissertation are as follows:

- A novel co-creative design system using AI models of conceptual similarity to support human ideation.
- A method for evaluating the effect of inspiration from a co-creative design system on design ideation.

- The impact of inspirational images selected for their conceptual similarity on design ideation during a specific design task.

AI-based co-creative design systems enable users to collaborate with an AI agent on open-ended creative tasks during the design process. This thesis presents a novel co-creative design tool that supports idea generation for new designs with two versions of the Collaborative Ideation Partner (CIP): CIP Sketch and CIP Design. The AI models for measuring similarity in the CIP use deep learning models as a latent space representation and similarity metrics for comparison to the user's sketch or design concept. The interactive experience allows the user to seek inspiration when desired. The concept of Collaborative Ideation Partner (CIP) provides a basis for developing other co-creative design systems and further exploration of design spaces of co-creative design systems.

Measuring ideation when co-creating with an AI-based co-creative design tool enables the comparison and evaluation of the impact of different AI models on the user's cognitive process and the creative outcome. This thesis presents an approach for measuring ideation that has two components: an aggregate analysis and a temporal analysis. The aggregate analysis adapts existing quantitative metrics for ideation: novelty, variety, and quantity of ideas expressed in the design session. The temporal analysis shows the temporal changes of novelty, variety, and quantity of ideas based on the AI contributions. These measures can be used in evaluating the impact of AI contributions in other co-creative systems that support design creativity.

This thesis emphasizes the effect of the AI-based inspirations based on the conceptual similarity. The findings from an exploratory study of the CIP Sketch and a user study of the CIP Design show that the AI model of conceptual similarity has a significant effect on the novelty, variety, and quantity of ideas during human design ideation. The implications of this study provide a basis for further exploration of the impact of AI-based inspiration on design ideation.

1.6 Thesis Overview

The structure of this thesis is as follows: Chapter 2 provides the background in computational co-creative systems and design ideation. In Chapter 3, the initial design of the Creative Ideation Partner (CIP Sketch) using visual and conceptual similarity is described. In Chapter 4, the pilot study that explores the effect of CIP contributions based on the visual and conceptual similarity is presented with an analysis, results from the data collected, and what we learned from the pilot study. In Chapter 5, the Creative Ideation Partner (CIP Design) is described to identify the components and the novelty of the approach to enhancing creativity with an AI model of conceptual similarity. In Chapter 6, the experimental study for comparing the effect of the AI model of conceptual similarity on ideation during a specific design task is presented with an analysis and results from the data collected. Finally, Chapter 7 provides a summary and discusses the limitations of the study, the implications for future research, and a plan for future research.

CHAPTER 2: BACKGROUND AND RELATED WORK

Two fields are germane to this thesis: computational co-creative systems and design ideation. This section will explore a number of computational co-creative systems and how they evaluate their systems and computational co-creativity. These projects contextualize this thesis in the broader field of computational co-creativity research. This chapter will also explain design ideation and measuring ideation introducing different methods and approaches. This can provide insights about how this study evaluates ideation in co-creative systems.

2.1 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. Co-creativity is a collaboration that multiple parties collaboratively and synthetically contribute to the creative process in a blended manner [62].

Computers can support human creativity in different ways. Lubart [55] classified four ways that computers can be involved in creative work: the management of creative work as a "nanny", communication between individuals collaborating on creative projects as a "pen-pal", the use of creativity enhancement techniques as a "coach", and the creative act through integrated human-computer cooperation during idea production as "colleague". Lubart [55] introduced the last category as the most ambitious vision of human-computer interaction for creativity that involves a real partnership. Co-creative systems present a computer as a colleague that intervenes in a creative process to generate, evaluate, or refine ideas. Co-creative systems also are associated with computational creative systems that autonomously generate creative products [14, 29] and creative support tools that support the users' creativity [15, 41]. Co-creative systems involve both characteristics of computational creative systems and creativity support tools [23, 49]. As a new type of computational creative system, co-creative systems themselves are creative generating some parts of an artifact. As a new type of creative support tool, co-creative systems contribute new ideas in a dialogue with humans providing inspiration. Depending on the collaboration dynamics of co-creative systems, the user can collaborate with the AI agent in a variety of ways.

Co-creative systems have been applied in different creative domains such as art, music, dance, drawing and game. Some of the co-creative systems directly perform actions to a shared artifact or performance whereas others provide suggestions to inspire users for generating novel ideas. This represents how a co-creative AI agent contributes to the creative process and can be a factor that distinguishes different co-creative systems. One of the co-creative interaction paradigms is an AI agent performing actions with a user simultaneously. Shimon [41] is a robotic marimba player that listens and responds to a musician in real time. This improvisational robotic musician performs accompaniment with the users' musical performance simultaneously. GenJam [4] is a jazz improvisation system that generates jazz improvisations detecting musical input. The system uses several jazz improvisation schemes to provide an accompaniment. These two examples in the music domain show the co-creative interaction that an AI agent performs actions with the user for a shared artifact simultaneously.

Another co-creative interaction paradigm is a turn-taking action between a user and an AI agent in a shared artifact. The examples for this interaction include art, design, and dance domains. Drawing Apprentice [21] 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 [68] is another co-creative drawing system, an AI interface that allows users and the AI agent to draw pictures collaboratively. DuetDraw helps users perform drawing tasks, such as completing the rest of the object that the user was drawing, drawing the same object in a different style, suggesting an object that matches the picture, finding an empty space on the canvas, and automatically colorizing the sketches. Cobbie [53] is a mobile robot embedded with recurrent neural network (RNN)-based co creative methods and mobile drawing system to support early-stage ideation. Cobbie provides inspirational sketches under the command of the designer. Viewpoints AI (VAI) is a co-creative dance partner that analyzes the user's dance gestures and provides complimentary dance in real-time by a virtual character projected on a large display screen [43, 44]. The user initiates dancing and the virtual character which is a life-sized silhouette performs dancing based on the user's dance gesture.

While the co-creative interaction paradigms above show the examples that an AI agent is directly involved in a creative activity as performing the same type of action with a user, another co-creative interaction paradigm is providing suggestions to the user. Sentient Sketchbook [86] and 3Buddy [56] are co-creative systems for game level design. In both systems, the AI agent provides feedback and additional ideas to develop the game design. These systems use a turn-taking interaction but provide suggestions to the human designer rather than creating game level directly.

A co-creative system that is closely related to our co-creative system is the Creative Sketching Partner [22, 50]. The Creative Sketching Partner is an AI-based co-creative sketching system that supports the conceptual design process. This AI partner presents sketches of varying visual and conceptual similarity based on the designer's sketch. The goal of the partner is to present a sketch to inspire the user to explore more of the design space and to reduce design fixation. Users can control the parameters of the algorithm by specifying how visually and conceptually similar to the system's sketch should be to their own. This study focuses on identifying the relationships between an AI model of conceptual shifts in a co-creative sketching system and three types of design creativity: combinatorial, exploratory, and transformational. The findings suggest that inspiration related to conceptual similarity is more associated with transformational creativity and inspiration related to visual similarity is more associated with combinatorial creativity. Our co-creative system is similar to the Creative Sketching Partner in terms of an AI model providing inspiring images based on conceptual similarity. However, this thesis focuses on the cognitive process of ideation during collaborating with the AI partner rather than the outcome of design (i.e. combinatorial, exploratory, and transformational creativity). While the Creative Sketching Partner focuses on how a co-creative sketching system can guide users towards different types of creative design, this thesis focuses on how conceptual similarity affects users' ideation differently.

2.2 Evaluating Co-Creative Systems and Computational Co-Creativity

Evaluating co-creative systems is still an open research question and there is no standard metric to measure computational co-creativity [49]. The research on co-creative systems shows various approaches to evaluate co-creative systems and computational co-creativity. Some researches focus on interactive experience and others focus on the effectiveness of the system in the evaluation. They also show different metrics and methods to evaluate the system. This section describes the examples of evaluation in computational co-creative systems.

The evaluation of Shimon [41], a robotic marimba player, is a performance-based evaluation of the system. The evaluation used observation to analyze the system's behaviors and the audience reactions during the performance. Drawing Apprentice [21] focused on usability and system accuracy in the evaluation. The evaluation methods include algorithm testing, voting, survey, and retrospective protocol analysis to evaluate the system and interactive experience. DuetDraw [68], a web-based co-creative drawing system, evaluated the user experience in collaboration with AI using a survey, think-aloud method, and semi-structured interviews. Cobbie [53], a mobile co-creative sketching robot, focused on user experience and co-creativity using self-rating questionnaires based on USE questionnaire [57] and Creative Support Index (CSI) [10], Semi-structured interview, Rating for ideation effectiveness (i.e. quality and novelty). In the evaluation of Viewpoints AI [43, 44], the researchers observed how participants interact with the systems and the participants provided feedback about their interactive experiences. While the examples of evaluation above focus on interactive experience in the creative process, Sentient Sketchbook [86], and 3Buddy [56] focused on the usefulness of the system since both systems support a goal-directed design rather than an open-ended artistic performance. They used a survey, interview, and observation to measure the usefulness of the system.

Karimi et al. [49] presented a framework for evaluating creativity in computational co-creative systems. They presented four main questions that can serve to characterize the many and varied approaches to evaluating computational models of co-creativity: Who is evaluating the creativity, what is being evaluated, when does evaluation occur and how the evaluation is performed. These questions provide a framework for comparing how existing co-creative systems evaluate creativity. According to the framework, our research is classified as Who: third party; When: formative & summative; How (metric): ideation effectiveness; How (method): experiment & retrospective protocol analysis; What: product & interactive experience. Specifically, the questions of How and What are the focus of our research. This thesis focuses on how inspirations that an AI model provides based on visual and conceptual similarity influence the users' ideation through evaluating their idea generated during creative design tasks.

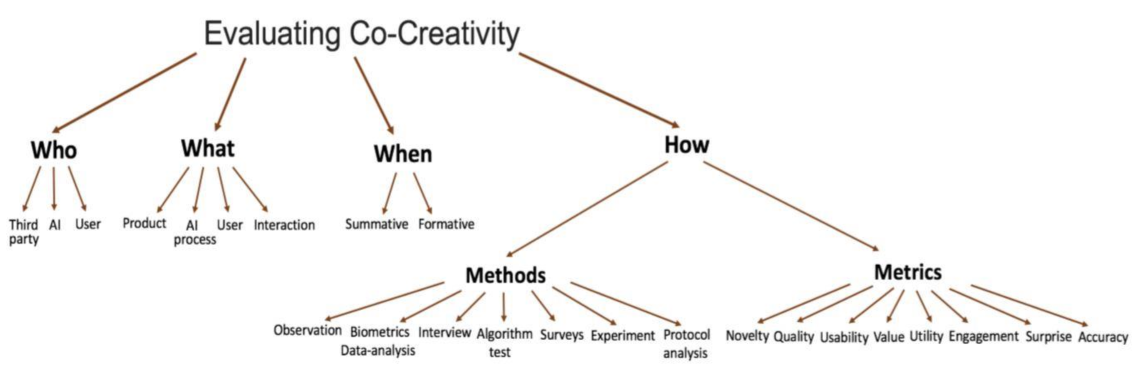


Figure 1: A hierarchical tree of evaluating creativity in computational co-creative systems [49].

The evaluations of existing co-creative systems described above, mostly focus on the interactive experience and the final product through evaluating the usability. However, they do not explore how the co-creative systems influence creativity in the creative process and what factors influence the user's creativity. This thesis focuses on understanding the cognitive processes in idea generation for evaluating the co-creative system through a process-based approach based on the cognitive process during the human-AI collaboration.

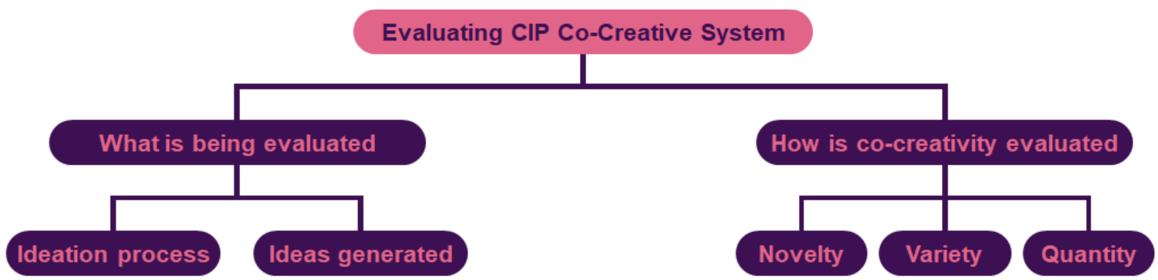


Figure 2: Evaluating the CIP design.

This thesis responds to "what is being evaluated" and "how is co-creativity evaluated" by evaluating the novelty, variety, and quantity of ideas in the ideation As shown Figure 2. Section 5 describes how we define and measure ideation in our co-creative system in more detail.

2.3 Design Ideation

This thesis claims that AI inspiration will affect design ideation. This thesis posits that the AI inspiration based on visual and conceptual similarity can influence the distinct properties of idea generation in the cognitive process. This thesis focuses on identifying the effect of AI inspiration on design ideation through evaluating the design ideation that involves the contribution of the AI partner in the idea generation.

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 [1, 3, 7, 17, 19, 54]. 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 [18, 67]. A design process is an evolution of different kinds of representations [38]. In a design process, designers exteriorize and visualize their design intentions and communicate with external visualizations to interact with their internal mental images [27]. During ideation, designers commonly use freehand sketches, and rough physical models as a tool for constructing external representations, cognitive artifacts of design [83]. Making sketches and physical models is an interaction and conversation [27]. In the ideation stage, designers frame problems producing new discoveries through the conversation. The graphical and physical representations, cognitive artifacts, thus are essential in the ideation process.

Many ideation methods have been developed to support designers as they generate innovative design solutions. Ideation methods provide a normative procedure on how to overcome certain blocks to creativity [40]. A comprehensive classification of ideation methods classifies the methods into two categories: intuitive and logical. Intuitive methods use mechanisms to overcome mental blocks (e.g. Brainstorming, Random Stimuli, Checklists, and C-Sketch) while logical methods involve systematic decomposition, analysis of the problem, databases, and physical principles (e.g. morphological charts and TRIZ). Analogy is an ideation method and we focus on analogy in this study. Analogical reasoning is an inference method in design cognition to develop a design leading to unexpected discoveries [36]. Design-by-Analogy (DbA) is a design tool that provides inspiration for innovative design solutions [12, 30, 37, 45]. Inspirations in Design-by-Analogy (DbA) are achieved by transferring a design problem (source) to a solution (target) in another domain [65]. The association between a source design and a target design can be either semantic (conceptual) characteristics or visual (structural) representations. The semantic and visual stimuli thus can be a basis for developing computational systems that support design ideation.

Collaborative ideation can support people in generating more creative ideas by exposing them to ideas different from their own. Analogy-based design positively impacts design ideation and creativity. However, the effect of analogical distance on design outcomes is controversial [45]. Some researchers argue that far-field analogies are more beneficial and others claim that near-field analogies are beneficial on ideation effectiveness [45]. Chiu and Shu [11] showed that far-field analogies increase the novelty and quality of ideation. Kennedy et al. [51] showed that far-field analogies are beneficial on ideation effectiveness for novelty, relevance, and effectiveness of solutions. On the other hand, Srinivasan et al. [78] showed that near-field analogies are used more frequently, and they found that far-field analogies increase novelty but decrease quality. Chan et al. [8] examined competing theoretical recommendations, Associationist theory, and SIAM theory, for how inspirational delivery systems on collaborative ideation platforms should account for semantic distance of inspirational stimuli. Associationist theory suggests that exposing ideators to ideas that are semantically far from their own maximizes novel combinations of ideas. On the other hand, SIAM theory cautions that systems should offer far ideas only when ideators reach an impasse, and offer near ideas during productive ideation, which maximizes exploration within categories. The results show that far inspirations can be harmful for creativity if delivered during productive ideation, and that collaborative inspiration systems could be improved by accounting for ideators' cognitive states.

This thesis posit that the AI inspiration based on conceptual similarity can influence the distinct properties of idea generation in the cognitive process. This thesis focuses on identifying the effect of AI inspiration on design ideation through evaluating the design ideation that involves the contribution of the AI partner in the idea generation. I consider conceptual similarity as a key factor for collaborative ideation using Design-by-Analogy

(DbA). This thesis investigates how the AI-based inspirations, conceptually similar images to a target design, impacts design ideation in a co-creative system.

2.4 Measuring Ideation

In this thesis, measuring ideation is a key to validating the claim that the AI inspiration based on conceptual similarity will influence design ideation in human-AI collaboration. The first step for measuring ideation is to define what an idea is in the ideation process using a co-creative system. Defining an idea in design ideation using a co-creative system is a challenge since the idea can be defined differently involving the contribution of an AI partner in ideation. In engineering design, an idea is normally considered as a possible solution to a given problem for evaluating the performance of idea generation [76]. However, an idea can be variously defined as a contribution that contains task-related information, a solution in the form of a verb–object combination, and a specific benefit or difficulty related to the task [74]. In order to identify what an idea is in the ideation using a co-creative system, this section will explore the metrics and methods for measuring ideation.

Evaluation of ideation methods can be classified into two groups: outcome-based approaches and process-based [66]. Outcome-based approaches focus on evaluating the ideation process based on the designs, or outcomes, and 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. This section describes possible metrics and methods for evaluating design ideation using a co-creative system.

2.4.1 Outcome-Based Approach

Outcome-based approaches have become more prevalent than process-based approaches due to the inherent complexity and difficulties in using process-based approaches [76].

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. Sarkar and Chakrabarti [75] considered novelty and usefulness as measures of creativity. Nelson et al. [66] proposed a simple metric that combines novelty and variety to measure the amount and quality of design space exploration. Maher et al. [60] and Grace et al. [39] employed novelty, value, and surprise to evaluate design creativity. Taylor et al. [79] considered quantity and subjective assessments of quality of the ideas as measures of ideation effectiveness. Shah et al. [76] introduced four types of outcome-based metrics for measuring ideation effectiveness, 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 the 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. Higher scores for novelty, variety, and quantity indicate greater exploration of the design space during the ideation. These metrics thus evaluate a designer's exploration and expansion of design space, existing ideation methods, and predict performance in various design activities. However, the measures and methodology do not consolidate the scores for all four measures into an overall effectiveness measure.

This thesis adapts existing quantitative metrics for ideation (i.e. novelty, variety, and quantity) to evaluate the effect of a co-creative design system on design ideation.

2.4.2 Process-Based Approach

Process-based approaches evaluate idea generation based on the cognitive processes via a protocol analysis and cognitive models. The Function-Behavior-Structure (FBS) ontology [32, 35] is a well-known methodology to analyze design activities by protocol analysis for understanding design cognition. The FBS ontology [32, 35] is a design ontology that describes all designed things, or artifacts, irrespective of the specific discipline of designing. The FBS ontology [32, 35] models designing in terms of three classes of ontological variables: function, behavior, and structure. 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 [33, 34]. As a consequence, there are 8 design processes that are numbered in Figure 3.

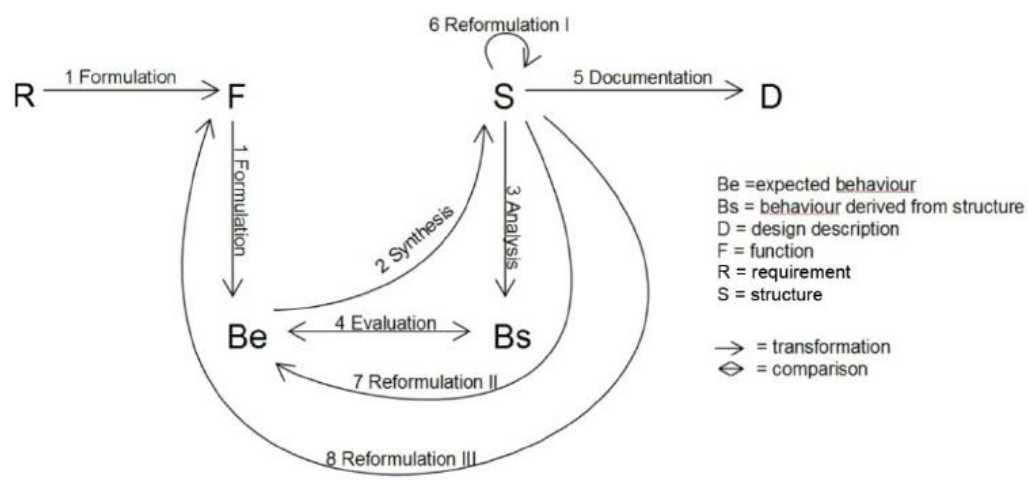


Figure 3: The FBS ontology [35].

The design process can be viewed as interactions between two notional design "spaces" of problem and solution [26, 59]. The Problem-Solution (P-S) index [34, 46] 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 [34, 46]. For the P-S division, design issues based on FBS are categorized into problem-focused and solution-focused design issues. P-S index reclassifies design issues and syntactic design processes into these two categories, as presented in Table 1.

Table 1: Mapping FBS design issues & processes onto problem and solution spaces [34]

Problem/solution space	Design issue	Syntactic Design Processes
Reasoning about Problem	Requirement (R) Function (F) Expected Behavior (Be)	1 Formulation 7 Reformulation II 8 Reformulation III
Reasoning about Solution	Behavior from Structure (Bs) Structure (S)	2 Synthesis 3 Analysis 4 Evaluation 6 Reformulation I

The P-S index helps to characterize the overall cognitive pattern of a design session indicating how the problem or solution focus is organized in the design cognitive process. 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. P-S index can be used for analyzing both a meta-level view (i.e., a single-value measurement) and a dynamic view (i.e., taking the sequential order of design issues/processes into consideration) to compare design cognition while using different creativity techniques for concept generation in collaborative engineering design settings.

$$\text{P-S index(design issue)} = \frac{\sum(\text{Problem}-\text{related issues})}{\sum(\text{Solution}-\text{related issues})} = \frac{\sum(R,F,Be)}{\sum(Bs,S)}$$

$$\text{P-S index(syntactic processes)} = \frac{\sum(\text{Problem}-\text{related syntactic processes})}{\sum(\text{Solution}-\text{related syntactic processes})} = \frac{\sum(1,7,8)}{\sum(2,3,4,6)}$$

Co-evolution is an interactive view on the problem-solving of design in a design process. In the problem-solving view of design, some researchers claim that the design process begins with an exploration within the problem space and others argue that design thinking is primarily solution-focused jumping into the solution space before the problem is formulated [46]. Maher et al. [61, 70] proposed to model this problem-design exploration as co-evolution. Co-evolution in design is an iterative process, designers re-interpreting a design problem based on the exploration of possible solutions until a good 'fit' between problem and solution [26, 59].

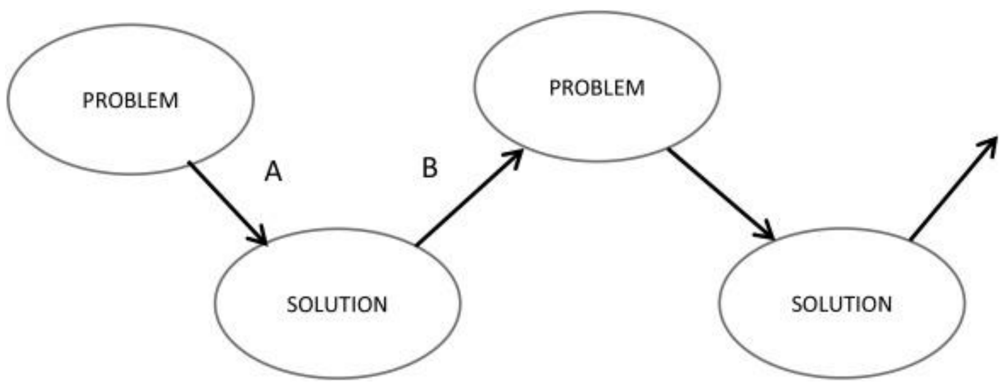


Figure 4: The co-evolution between problem and solution [25]

Figure 4 shows co-evolution as a series of unarticulated 'jumps' that bridge the gap between the problem space and solution space. There are two types of transitions in co-evolution: 'downward jump' and 'upward jump'. A downward jump is a transition from problem space to solution space. An upward jump is a transition from solution space to problem space. The transitions in co-evolution, a 'jump' between the problem space and solution space, can be used to understand a design process and designers' cognition reflecting complex problem situations the designers are facing. The sequential process of co-evolution is a descriptive framework for designing in both problem solving and reflective practice paradigms.

CHAPTER 3: INITIAL DESIGN OF COLLABORATIVE IDEATION PARTNER (CIP SKETCH)

3.1 System Overview

Collaborative Ideation Partner (CIP) is a co-creative sketching tool which builds on previous works [22, 50], that interprets sketches drawn by a user and provides inspirational sketches based on visual similarity and conceptual similarity. The CIP was developed to explore the effect of an AI model for visual and conceptual similarity on design ideation in a co-creative design tool.

The interface of CIP is shown in Figure 5. 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. 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 provides different categories of ideation stimuli.

Figure 5 shows an example of an inspiring sketch and how participants communicate with an inspiring sketch to develop their design. The design task shown in Figure 5 is to design a chair for a gaming computer desk. The participant drew a basic chair with back, seat, legs, and small wheels before requesting inspiration from the AI partner. The sketch suggested from the AI models is a bulldozer: visually similar and conceptually different to the participant's sketch. After getting the inspiring sketch, the participant made the wheels much bigger for better mobility and added a leg rest for comfort. During the retrospective protocol, the participant described that *"I decided to go with bigger wheels here, just thinking of bulldozer, little more heavy duty. I mean, I also noticed the little lift gate or whatever that is. And that kind of made me think that I needed to add like some kind of leg support and that kind of made sense."*.

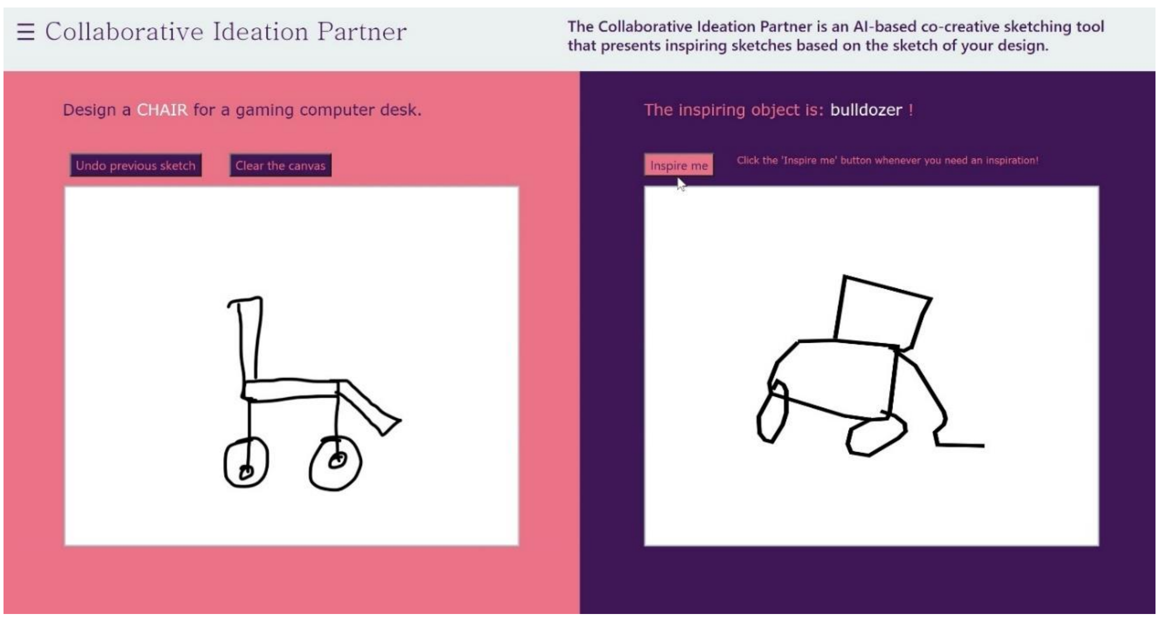


Figure 5: User Interface of Collaborative Ideation Partner Sketch

The components of the CIP system include the sketch database, word embedding models, an AI model for visual similarity, an AI model for conceptual similarity, and the front-end user interface. The system diagram shown in Figure 6 shows how the system generates an inspiring sketch. The turn-taking interaction between the user and the AI partner is designed to facilitate communication for design ideation. Once the user draws a sketch as input then clicks the "inspire me" button to get the system's response, the system's turn will be a sketch of another object with the corresponding label shown on the inspiring sketch space (purple area). To generate an inspiring sketch, the AI model of visual similarity computes the visual similarity based on the vector representations of visual features of the sketches using the QuickDraw sketch dataset and the AI model of conceptual similarity calculates the conceptual similarity based on the category names of the sketches using two pre-trained word2vec models. The system is designed to support the exploratory study and provides a different set of inspiring sketches in different categories of ideation stimuli, for each design task. The design tasks and different categories of ideation stimuli are:

- Sink: Random sketches
- Bed: Conceptually and visually similar
- Table: Conceptually similar but visually different
- Chair: Visually similar but conceptually different

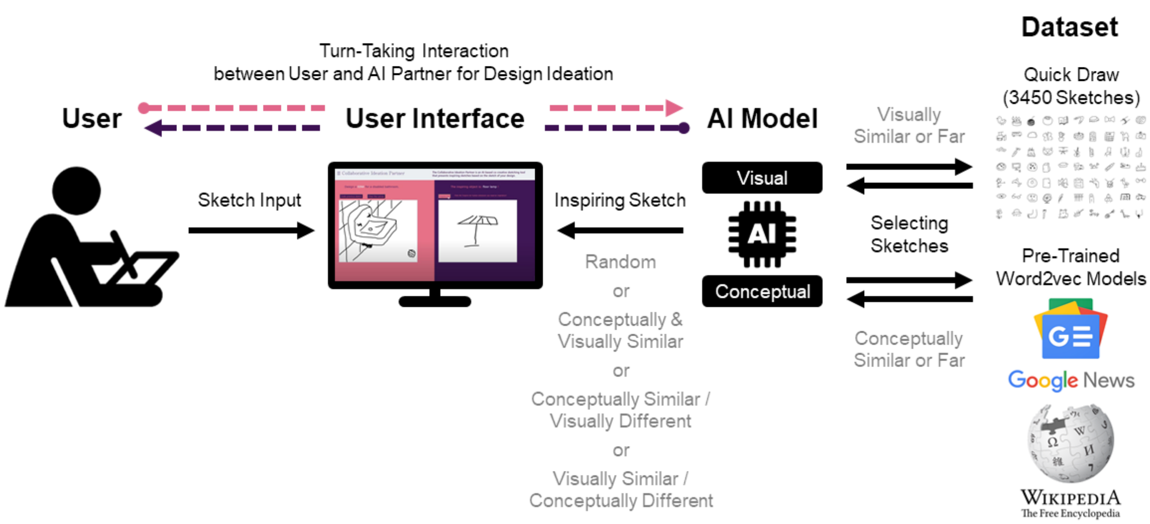


Figure 6: System Diagram of Collaborative Ideation Partner Sketch

3.2 Sketch Dataset

For the source of inspiring sketches, the CIP system uses a public benchmark called QuickDraw! [47], 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 [28, 71] 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.

3.3 AI Models for Visual Similarity 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. The visual similarity component selects sketches from the sketch dataset based on a representation of the stroke data in the image file. The conceptual similarity component computes the degree of similarity between the category names of the objects in design tasks and the category names in the objects in the sketch dataset. For example, "bed" and "pillow" are more likely to appear in the same context and potentially associated to each other compared to "bed" and "animal migration" which are less likely to appear together.

For measuring visual similarity, we train a model with 3 convolutional layers, 2 LSTM layers, and a softmax output layer on the QuickDraw dataset. For all the sketches in the dataset, we collect the last LSTM layer of the trained model and use that as the vector representations of visual features of the sketches. Because of the diverse nature of the sketches in the same category, we identified 10 clusters of sketches for all the categories using the elbow method. We cluster the sketches using the K-means algorithm and randomly select one sketch of each cluster as a typical sketch for that cluster of sketches. Thus, we convert the QuickDraw dataset of 50 million sketches into 3450 sketches (345 categories, each has 10 sketches). CIP collects user sketches as the array of x and y coordinates of strokes and simplifies the strokes using Ramer–Douglas–Peucker algorithm. 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. CIP takes the last LSTM layer of the trained model as the vector representation of visual features of the user sketch. We calculate the Euclidean distance score of visual features of user sketch. A high score means sketches are more visually similar and low scores means sketches are less visually similar. We prepare a sorted list of sketches to generate the sequence of sketches in the conceptual similarity model.

For measuring conceptual similarity, 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 [63] and Wikipedia [73], 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.

We use these two AI-based components of the CIP to generate sequences of sketches with combinations of visual and conceptual similarity to the user's current sketch and design task to inspire the user during their design process and measure the effect of visual and conceptual similarities on ideation.

3.4 User Experience: 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. One of the modes (i.e. sink) uses a random sketch selection while three other modes use AI models to select an inspiring sketch as inspiration in CIP.

- **Random**: 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.
- **Similar**: 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). To generate the set of inspiring sketches, the model selects the top 15 conceptually similar objects to the concept of a bed. For each concept, the model selects one sketch from each of the 10 clusters of sketches associated with that concept and thus 150 sketches (15 X 10) are selected in this phase. The model then selects the sketch that is most visually similar to the participant's sketch from the 150 sketches to provide an inspiring sketch.
- **Conceptually similar**: 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). To generate the set of inspiring sketches, the model selects the top 15 conceptually similar objects to the concept of table in the dataset. For each concept, the model selects one sketch from each of the 10 clusters of sketches associated with that concept and thus 150 sketches (15 X 10)

are selected in this phase. The model then selects the most visually dis-similar sketch to the participant's sketch from the 150 sketches to provide an inspiring sketch.

- **Visually similar**: 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). To generate the set of inspiring sketches, the model selects the top 150 visually similar objects that do not contain any sketches of a Chair in the dataset. The model then selects the most conceptually dis-similar object to the concept of a chair from the 150 ranked dissimilar sketches to provide an inspiring sketch.

CHAPTER 4: EXPLORATORY STUDY OF DESIGN IDEATION WITH THE CIP SKETCH

This section presents a study to explore the effect of an AI model of visual and conceptual similarity as the basis for inspiration during a design process. The goal of the exploratory study is to explore the effect of AI inspiration on ideation through an analysis of the correlation between conceptual and visual similarity with characteristics of ideation. Specifically, we are interested in the relationship between the users' ideation and sources of AI inspiration. This exploratory study focuses on identifying distinct patterns of the participant's ideation in a human-AI collaboration where the AI partner contributes content based on visual and conceptual similarity.

4.1 Pilot Study Design

The type of study is a mixed design of between-subject and within-subject design. 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). The control condition (condition A) for each group is the same but the treatment condition for each group is different (condition B or C or D). The control condition and 3 treatment conditions are the different types of inspirations presented in Section 3.4:

- 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 pilot study, we collected video protocol data of design sessions and retrospective protocol sessions. 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 CCI 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. This study intended to recruit 30 participants (10 participants for each participant group). However, due to COVID 19, it was difficult to recruit participants and we decided to go ahead with the data from the 12 participants as an exploratory study. The results from 12 participants are intended to guide the research to be done after the pilot study rather than to provide results.

The task is an open-end design task in which participants were asked to design an object in a given context through sketching. Different objects for the design task were used for each condition: a sink for an 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, 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 as shown in Table 2. 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 while watching their design session recording as time goes on, and how the AI's sketches inspired their design in the retrospective protocol session.

Table 2: Counterbalanced conditions of 3 within-subject groups for the design tasks

A&B			**A&C**			**A&D**		
	Task 1	Task 2		Task 1	Task 2		Task 2	Task 2
P1	A	B	P2	A	C	P3	A	D
P4	B	A	P6	C	A	P5	D	A
P7	A	B	P8	A	C	P9	A	D
P10	B	A	P11	C	A	P12	D	A

4.2 Analysis of Data Collected

To measure ideation in a co-creative system, we developed four metrics based on [76], used for evaluating idea generation in design: novelty, variety, quality, and quantity of design. The four metrics of the outcome-based approach can be a basis for identifying specific properties associated with the visual and conceptual similarity in design ideation and the process-based approach can help us understand the cognitive process in the design ideation collaborating with the AI partner. This study applies the basic principles of both outcome-based and process-based approaches but redefines and expands the specific measures for a co-creative system.

4.2.1 Data Collected

To evaluate the impact of differences in visual and conceptual similarity of the AI partner's contribution to ideation during the use of CIP, this study used both an outcome-based approach and a process-based approach. Two types of data were collected for analyzing the study results: a set of sketches that participants produced during the design tasks and the verbalization of 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.

4.2.2 Data Segmentation and Coding

To analyze the verbal data collected from the retrospective sessions, we adapted the FBS coding scheme for characterizing cognitive issues during a design process [32, 35]. An idea can be variously defined as a contribution that contains task-related information, a solution in the form of a verb–object combination, and a specific benefit or difficulty related to the task [74]. The FBS coding scheme provides a segmentation into individual ideas associated with specific cognitive issues in design. First, the verbal data of all retrospective protocol sessions 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 [32, 35] as an idea segment, since an inspiring sketch segment includes multiple ideas. The idea segments were coded based on FBS ontology [32, 35] as requirement (R), function (F), expected behavior (Be), behavior from structure (Bs), and structure (S). A segment coded R is an utterance that talks about the given requirement in the statement of design task (e.g.

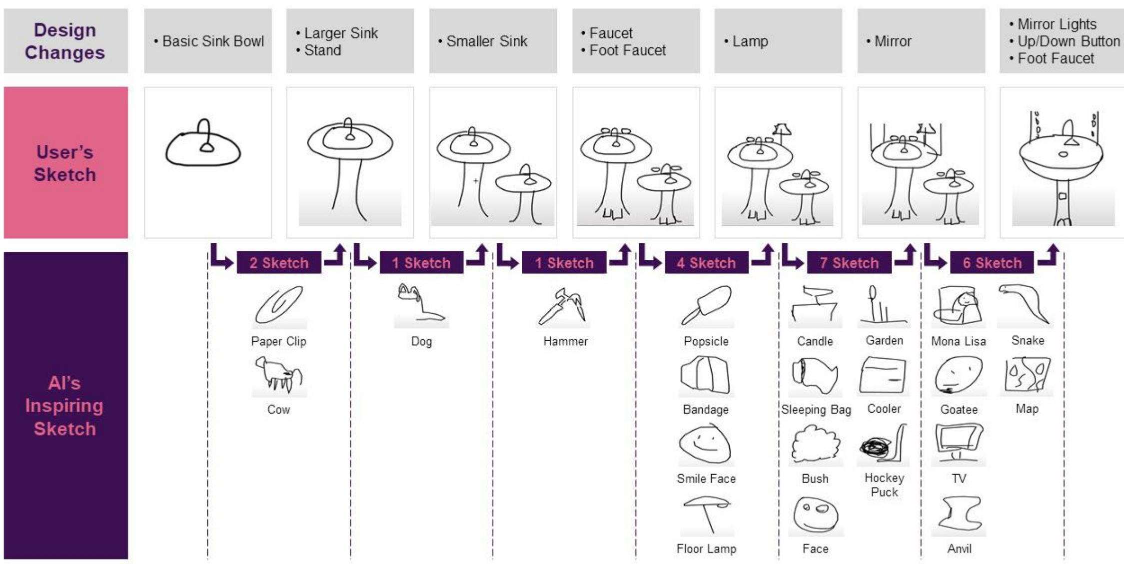
accessible bathroom); a segment coded F is an utterance that talks about a purpose or a function of the design object (e.g. more accessible); a segment coded Be is an utterance that talks about an expected behaviors from the structure (e.g. water could automatically come out); a segment coded Bs is an utterance that talks about a behavior derived from the structure (e.g. pressing on); a segment coded S is an utterance that talks about a component of the design object (e.g. button). The result of this coding scheme is a segmentation of the verbal protocol into individual ideas, each associated with one code: R, F, Be, Bs, S.

Two coders coded the idea segments individually based on the coding scheme above then came to consensus for the different coding results. The coding instruction was given to the coders included how to segment inspiring sketch segments and idea segments, how to code each idea segment with the coding scheme, and how to code new and repeated ideas. The two coders coded a design session together to make an initial agreement for segmentation and coding before coding individually then coded all design sessions individually. Once each coder completed coding all data individually, the two coders discussed each of the different coding results and came to consensus. The inter-rater reliability (IRR) was calculated to ensure the consistency of the coding between the two coders using the formula described in Miles and Huberman [64] and the IRR was 0.96. This indicates sufficient agreement among multiple coders that Miles and Huberman [64] suggested.

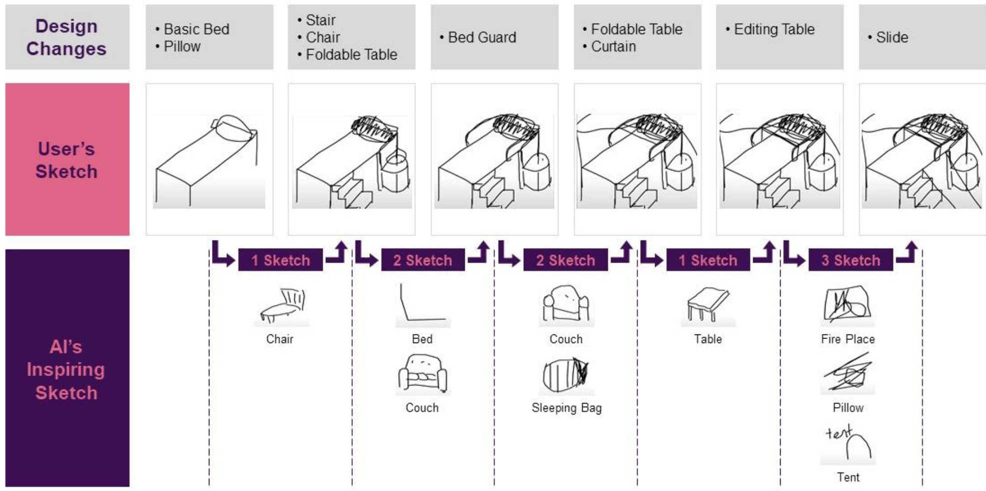
4.2.3 Observations of Collected Data

This section describes the examples of co-evolving design produced from participants, participants’ responses to inspirations on the use of CIP, a preliminary analysis of the coded data. To identify the patterns of the co-evolution of the participant's sketch and the AI inspired sketch, we observed the video stream data in addition to coding the transcript data. The examples of co-evolving design show 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. We also report the results of coded data.

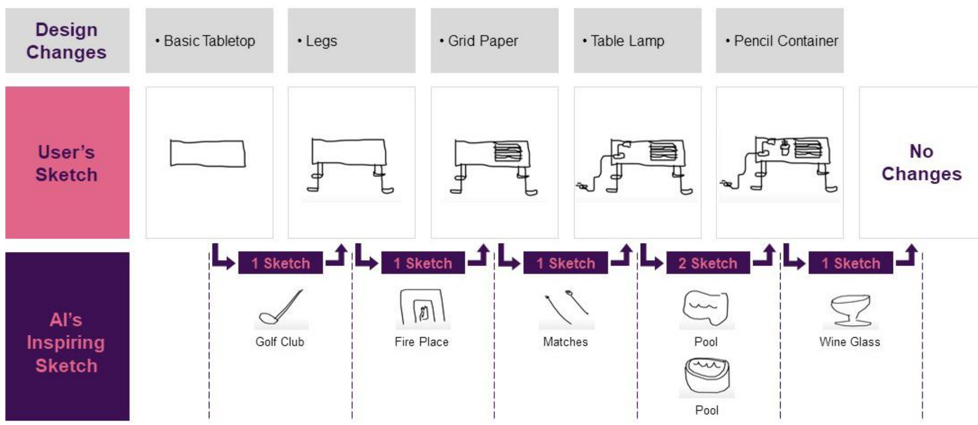
Figure 7 shows an example of the process for the evolution of the participant's sketch using CIP in each condition. In an evolution of the participant's sketch, participants in each condition start to design a basic shape of the target design then develop the design communicating with the AI partner. However, the participant in condition A shows a different pattern with condition B, C, and D. The participant in condition A explored many inspiring sketches but did not have many design changes while the participants in condition B, C, and D incrementally developed their design with inspiring sketches. To be specific, in Figure 7a, the participant drew a basic sink with a handrail before getting the first inspiration then tried to get an inspiration from the AI partner. The participant had 7 inspiring sketches but did not change anything for the design. The participant then cleaned all the canvas then drew a new sketch which is a sink with a motion sensor. The participant had 4 inspiring sketches and did not change anything again for the design. The participant cleaned the canvas and drew a new sketch again applying the motion sensor idea again then had 2 inspiring sketches. However, the participant finally finished the design without any changes. This case shows an example that participants do not have many ideas from random inspirations. On the other hand, the participants in conditions B, C, and D show a similar pattern of evolution applying some ideas from the inspiring sketches. For example, in Figure 7d, the participant drew a basic chair without any special function for the context of gaming before getting the first inspiration then tried to get an inspiration from the AI partner. The first inspiring sketch was a raccoon and the participant 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). After that, the participant had the second inspiring sketch which is a power outlet. The participant said the power cord gave a lot of ideas to develop the design and added a speaker on the ear decoration, buttons on the armrest to control sound volume/massage/lights, and power cord. In this case, the idea came from the inspiration sketch was transferred to new functions of the chair while the idea came from the raccoon was transferred to the shape of the chair. The participant then had 5 more inspiring sketches (i.e. rain, hurricane, zigzag, and camouflage). The participant mentioned that the participant was 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, the participant had nine more inspiring sketches, but did not change anything for the chair design. The participant mentioned got a decoration idea from the star shape of an aircraft carrier and snowflake but did not change the chair design.



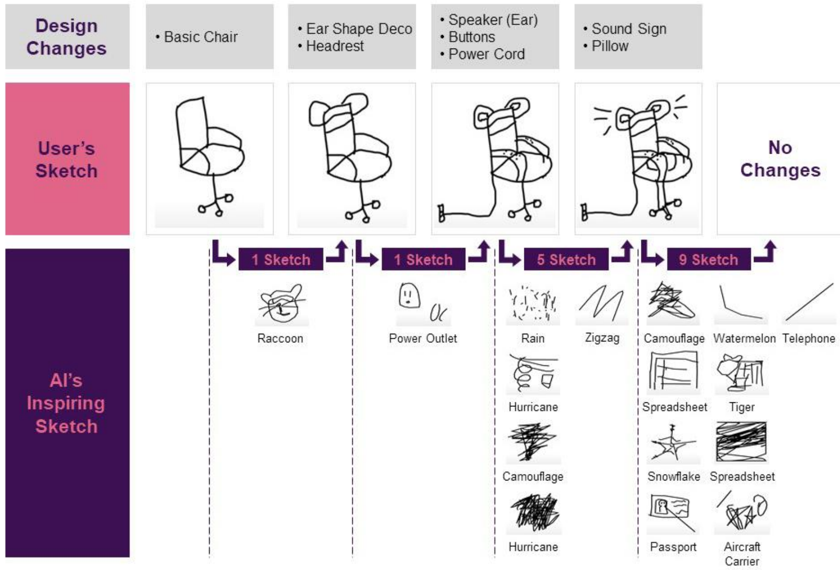
(a) An evolution of the participant's sketch in condition A (randomly)



(b) An evolution of the participant's sketch in condition B (visually and conceptually similar)



(c) An evolution of the participant's sketch in condition C (conceptually similar and visually different)



(d) An evolution of the participant's sketch in condition D (visually similar and conceptually different)

Figure 7: Evolution of participants' sketches using CIP sketch in each condition

From the observations of the video stream data, we found some distinct patterns of user responses to inspiration on the use of CIP. A common case is that a participant starts with a basic shape of the target design before getting the first inspiration then tries to get ideas from the AI partner exploring inspiring sketches as shown in Figure 7. In this case, when the participant gets an inspiring sketch from the AI partner, the participant tries to find an idea from the inspiration evaluating the inspiring sketch. The participant evaluates the relationship between the inspiring sketch and the target design to find an idea from the inspiring sketch. In this evaluating phase, some participants actively tried to use inspiring sketches for ideation while others did not try to use the inspiring sketches. For the cases that did not try to use the inspiring sketches, we assume that they did not find a direct relationship between an inspiring sketch and the target design or they wanted to develop their own idea rather than collaborating with the AI partner. Another interesting case different from the common case on the use of CIP is that a few participants produced a lot of ideas and developed their sketch before getting the first inspiration from the AI partner. The participants produced their own ideas based on the given context of the design task (e.g. disabled bathroom) and made a complicated sketch including their own ideas. In this case, the participant tried to interact with the AI partner once they completed their own initial ideation. Since the participants already developed the target design a lot, they used the inspiring sketches for finding additional features that can be attached to the design that they already created or looking for what they are missing in the design they already created evaluating the inspiring sketches. To find different patterns of user responses to inspiration on the use of CIP between the conditions, more data needs to be collected to learn.

4.2.4 Interpretation of Coded Data: Outcome-Based Approach

To identify specific properties of ideation associated with visual and conceptual similarity of the AI partner's contribution in a co-creative system, we applied both evaluation methods of ideation: an outcome-based approach and a process-based approach for the coded data. For the outcome-based approach, we measured four metrics commonly used for evaluating ideation: novelty, variety, quality, and quantity of design [76]. However, we redefined the specific criteria to apply the four metrics to our study.

The transcripts data were coded for 24 sessions of retrospective protocol (i.e. 12 sessions of condition A, 4 sessions of condition B, 4 sessions of condition C, and 4 sessions of condition D) as described in the section 4.2.2. This is a preliminary analysis of the collected data to move forward to the next step of analysis. In this analysis, we focused on identifying the number of ideas and the patterns of ideas generated in each condition rather than measuring effectiveness of ideation. A segment of R/F/B/S is defined as an idea and counted the number of ideas in each condition.

Novelty. Novelty is a measure of how unusual or unexpected an idea is as compared to other ideas (Shah et al., 2003). 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.

The results showed that all treatment conditions (B, C, D) have more novel ideas than the control condition (A) in the total number of novel ideas. Specifically, 10 participants out of 12 participants produced more novel ideas in a treatment condition than the control condition. When comparing the novelty of 3 groups, the group A&C showed the largest difference between the control condition and the treatment condition where condition C selected inspiring sketches that are conceptually similar and visually different. As shown Figure 9, all participants in the group of A&C produced more novel ideas in the condition C than the condition A while one of the participants (i.e. P4) in the group A&B and one of participants (i.e. P9) in the group A&D produced fewer novel ideas in the treatment condition than the control condition as shown in Figure 8 and 10. This result can indicate that the conceptual similarity of inspiring sketches may be associated with the novelty of ideas in the ideation with CIP sketch.

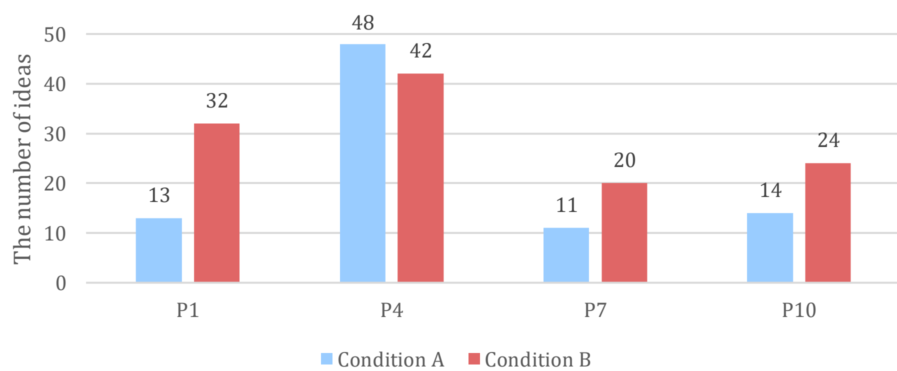


Figure 8: Novelty of ideas in the group of A&B

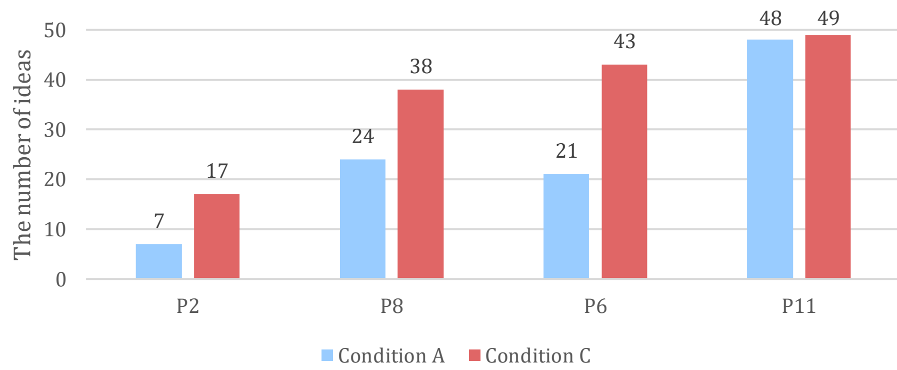


Figure 9: Novelty of ideas in the group of A&C

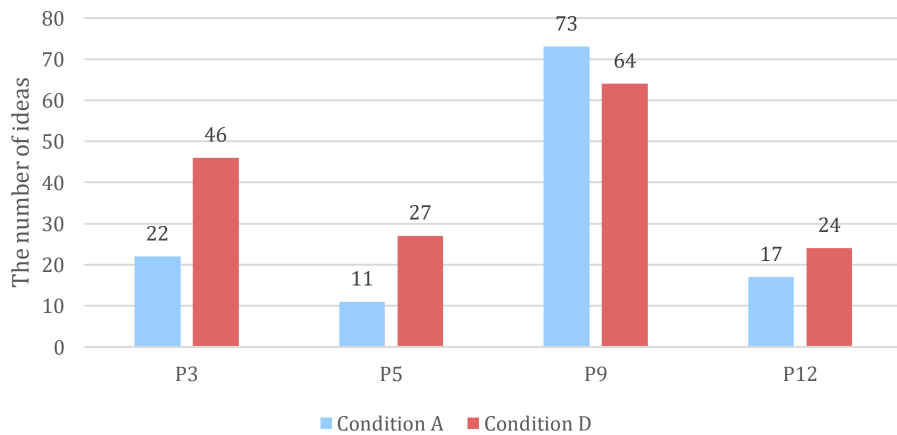


Figure 10: Novelty of ideas in the group of A&D

Variety. Variety is a measure of the explored solution space during the idea generation process [76]. For measuring variety in this study, each idea segment was coded whether it is a new idea or a repeated idea in a design session and 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.

Figure 11, 12, and 13 show the aggregated results of codes comparing the control condition (A) and one of the treatment conditions (B or C or D). The results show that all treatment conditions (B, C, D) show more new ideas than the control condition in the total number of new ideas. Specifically, 9 participants out of 12 participants produced more new ideas in a treatment condition than the control condition while one participant (P4) in condition A and one participant (P9) in condition D produced fewer new ideas in a treatment condition than the control condition. R (requirement) is the most prevalent idea in all conditions, and each treatment condition has a distinct pattern of codes when compared to the control condition. The pilot study does not have enough participants to show statistically significant differences. However, the results of the preliminary analysis can indicate that the treatment conditions may positively affect the ideation than the control condition and condition C may be associated with a variety of ideas in the ideation with CIP.

Figure 11 shows the number of new ideas in each code and the total number of ideas in the group of the condition A&B including four participants (i.e. P1, P4, P7, P10). The result of group A&B did not show a distinct pattern or a significant difference in each code between the participants. Although all participants in condition B produced more requirements than in condition A, only P1 showed a large difference between the condition A and B. The number of ideas in Function, Behavior, and Structure is similar between the condition A and B.

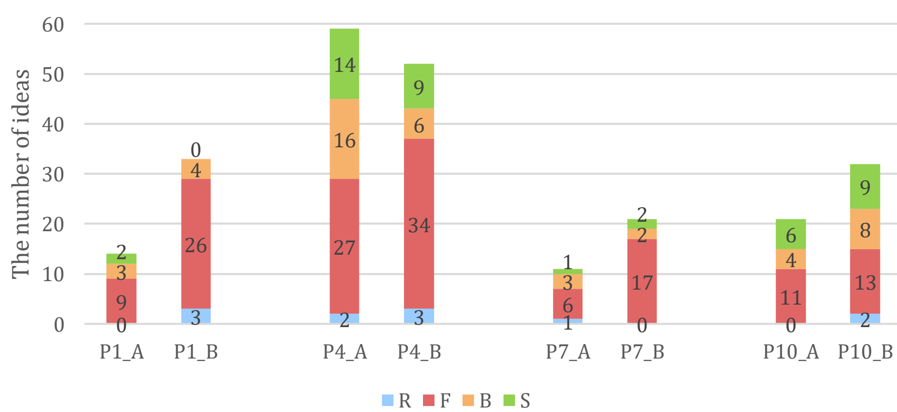


Figure 11: Variety of ideas in the group of A&B

While the group of A&B did not show a distinct pattern, the variety of ideas in condition C is higher than in condition A. Figure 12 shows the results of codes comparing the control condition (A) and one of the treatment conditions (C). The results of the group A&C show some distinct patterns in function. All participants produced more functions in condition C than in condition A. The number of function ideas showed a large difference for all participants between condition A and C. This result indicates that the conceptual similarity inspired the participants to produce more various functions associated with the context of the design.

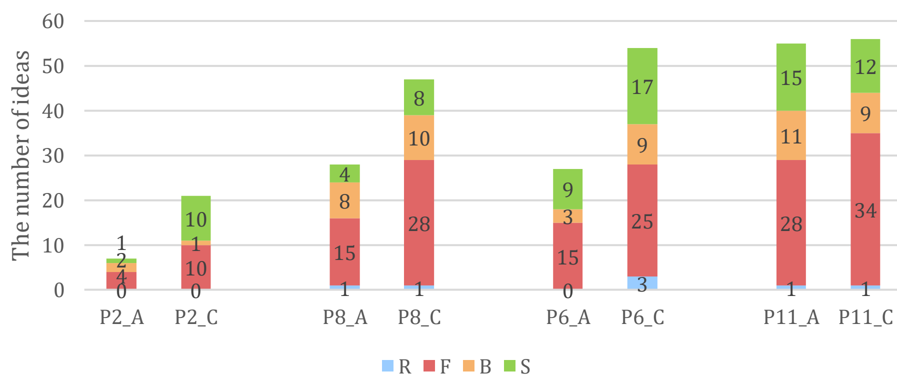


Figure 12: Variety of ideas in the group of A&C

While the group of A&C showed a large difference in Function, the group of A&D did not show a distinct pattern in the variety of ideas.

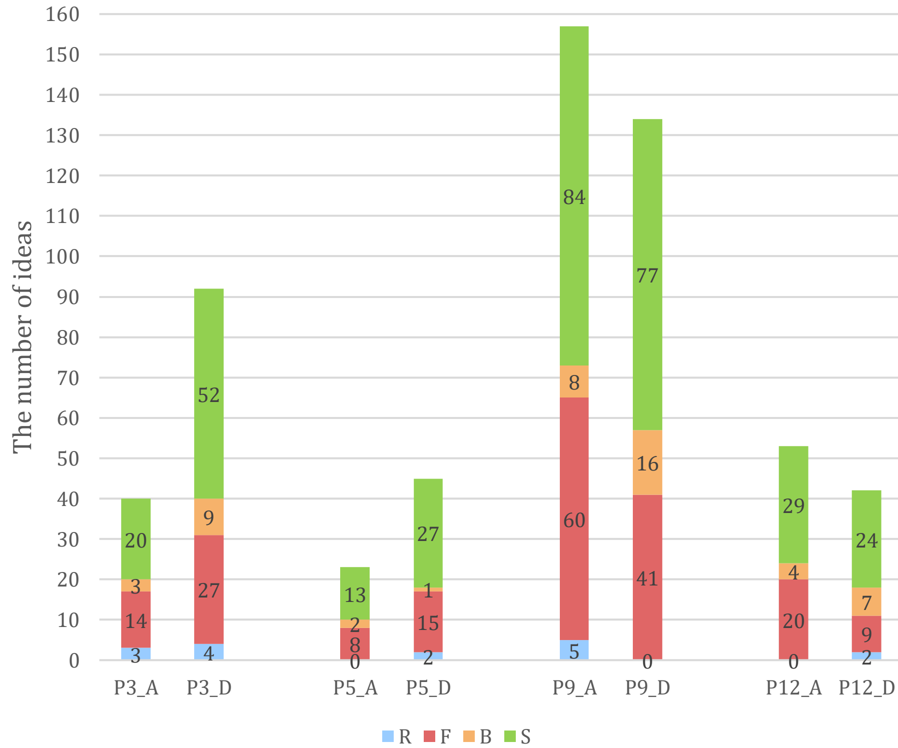


Figure 13: Variety of ideas in the group of A&D

Quality. Quality is a subjective measure of the design [76]. In this study, quality is measured using the Consensual Assessment Technique (CAT) [2], 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.

Tables 3, 4, and 5 show the result of the quality evaluation that each judge made for each design. The results show that most rates the judges rated for each design are consistent. To indicate the differences between the conditions, the table uses color codes, low-yellow; medium-orange; high-red.

The result of group A&B did not show a distinct pattern between the condition A and B. P1 showed the same quality between the condition A and B; P4 showed a higher quality in condition A than in condition B; P7 showed a higher quality in condition B than in condition A; P10 showed the same quality between the condition A and B.

Table 3: Quality evaluation results of each judge in the group of A&B

	Condition A		Condition B	
	Judge 1	Judge 2	Judge 1	Judge 2
P1	low	low	low	low
P4	high	high	medium	medium
P7	low	low	high	high
P10	medium	medium	medium	medium

While the group A&B did show a difference between condition A and B, two out of four participants in the group A&C showed higher quality in condition C than condition A. P2 and P8 produced higher quality in condition C than in condition A (i.e. P2: low to medium, P8: medium to high). The other two participants (i.e. P6, P11) produced high quality in both conditions A and C.

Table 4: Quality evaluation results of each judge in the group of A&C

	Condition A		Condition C	
	Judge 1	Judge 2	Judge 1	Judge 2
P2	low	low	medium	medium

P6	high	high	high	high
P8	medium	medium	high	high
P11	high	high	high	high

The results showed that the quality of ideas in condition D is higher than in condition A, where condition D selects sketches that are visually similar and conceptually different for inspiration. Table 5 shows the result of the quality evaluation that each judge made for each design in condition A and condition D. Three out of four participants produced higher quality in condition D than condition A. P3 produced much higher quality in condition D than condition A (i.e. low to high). P5 and P9 produced higher quality in condition D than in condition A (i.e. P5: low to medium, P9: medium to high). This result indicates that the visual similarity of inspiring sketches may be associated with the quality of ideas in the ideation with CIP.

Table 5: Quality evaluation results of each judge in the group of A&D

	Condition A		Condition D	
	Judge 1	Judge 2	Judge 1	Judge 2
P3	low	low	high	high
P5	low	low	medium	medium
P9	medium	medium	high	high
P12	low	low	low	low

Quantity. Quantity is the total number of ideas generated [76]. This study measures idea based on cognitive issues and every time an idea came up as a cognitive issue, it activates the brain about the idea. In this study, the quantity of ideas is defined as the total number of times any idea that is activated in cognitive issues. For measuring quantity, the number of ideas both new ideas and repeated ideas coded as R/F/B/S is counted in a design session. The results show a similar pattern to the result of variety with some distinct patterns.

In Figure 14, the results of the group A&B did not show a consistent result between the participants that can indicate a distinct pattern between condition A and B. For example, for the total number of ideas, two participants generated more ideas in condition B than in condition A but the other two participants generated more ideas in condition A than in condition B. Another example is the result of F (function). Although all four participants generated more functions in condition B than in condition A, the differences are not huge except P1. This result suggests further analysis to identify distinct patterns between the condition A and B, and relationships between the condition B and C and D since condition B has both characteristics of condition C and D.

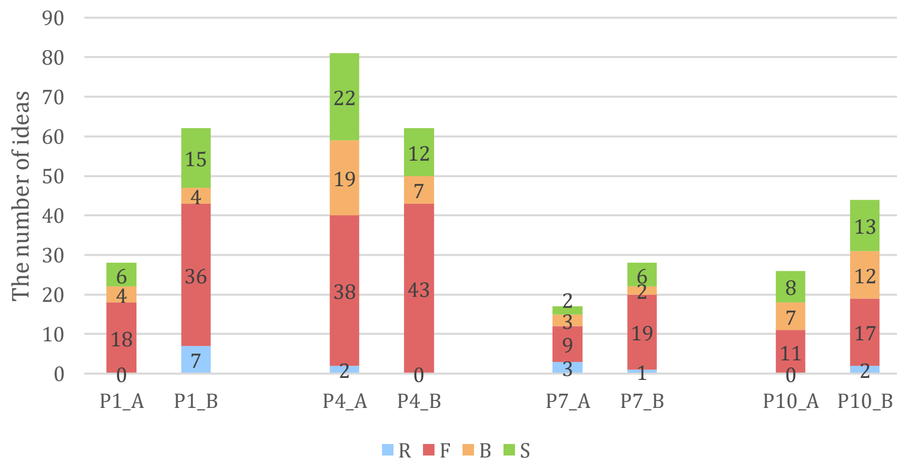


Figure 14: Quantity of ideas in the group of A&B

Figure 15 shows the results of the quantity of ideas in the group of A&C. The results show a similar pattern to the result of variety with some distinct patterns. First, for the total number of ideas, 3 out of 4 participants (i.e. P2, P8, P6) generated more ideas in condition C than in condition A. Second, 3 out of 4 participants (i.e. P2, P8, P6) generated more ideas of F (function) and S (structure) in condition C than in condition A. This result indicates that the conceptual similarity of inspiring sketches facilitates producing new functions and the emerging functions were transferred to structures of the design.

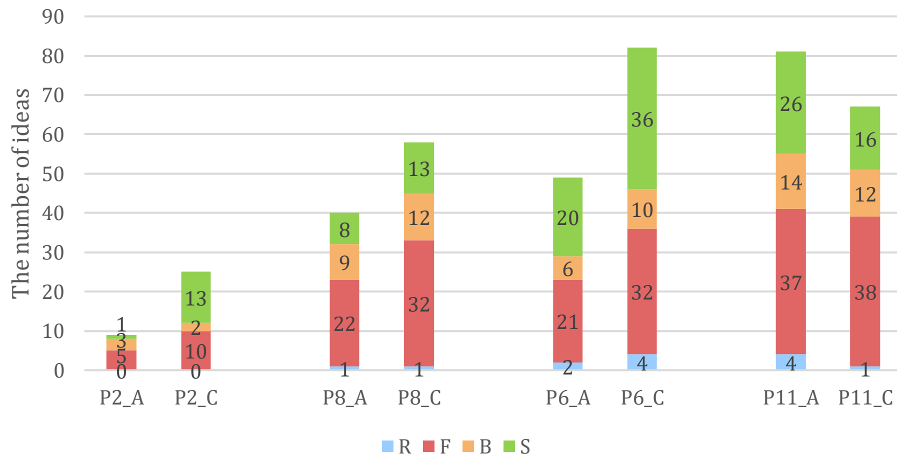


Figure15: Quantity of ideas in the group of A&C

Figure 16 shows the results of the group A&D did not show a consistent result between the participants that can indicate a distinct pattern between condition A and D. This result suggests that it needs further analysis to identify the effect of visual similarity associated with the functions.

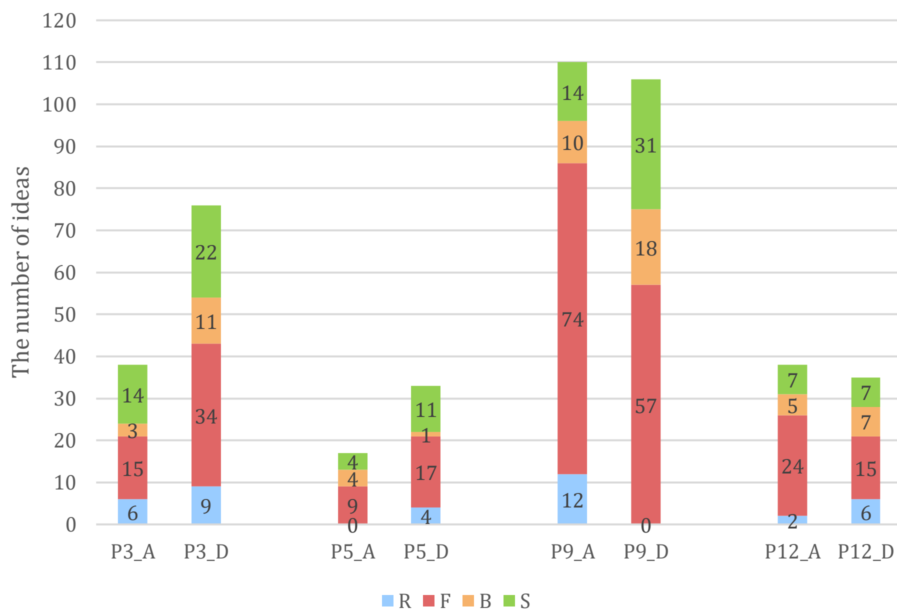


Figure 16: Quantity of ideas in the group of A&D

4.2.5 Interpretation of Coded Data: Process-Based Approach

For the process-based approach, this study used P-S index, integration of the FBS coding scheme with a Problem-Solution (P-S) division [46]. We examined the design cognition from a meta-level view (i.e., a single-value measurement) in this pilot study. For the meta-level view, the P-S index is calculated by computing the number of the total occurrences of the design issues/processes 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 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 indexes can facilitate comparisons across multiple sessions and across sessions involving different conditions in an effective way. Table 6 shows the P-S indexes for all design sessions. The results of the P-S index did not show a distinct pattern between the conditions. 23 design sessions out of 24 design sessions revealed a problem-focused style and only 1 design session (i.e. P2 in condition

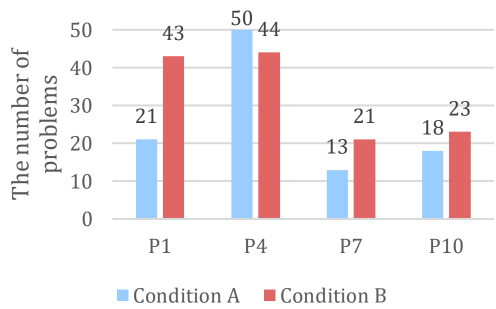
C) revealed a solution-focused style. From this result, we assume that the inspirations that the AI model are associated with the problem space rather than the solution space.

Table 6: P-S index for each design session

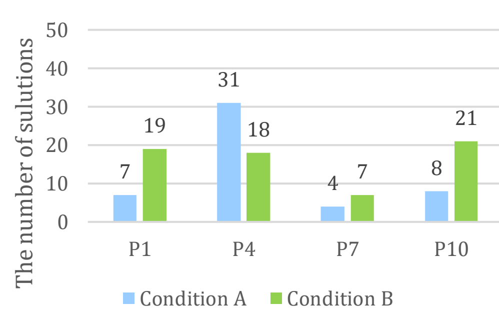
P-S Index (A&B)			P-S Index (A&C)			P-S Index (A&D)		
	A	B		A	C		A	D
P1	3.00	2.26	P2	2.00	0.67	P3	1.53	1.81
P4	1.61	2.44	P6	1.04	1.05	P5	1.83	2.00
P7	3.25	3.00	P8	4.00	2.05	P9	5.11	1.72
P10	2.25	1.10	P11	1.53	1.91	P12	3.75	2.50

Since the P-S index did not show a distinct difference between the conditions in a meta-level view of the design cognition, this study focused on the differences between the number of problems and solutions to identify the effect of visual and conceptual similarity. Figures 17, 18, and 19 show the aggregated results of codes comparing the control condition (A) and one of the treatment conditions (B or C or D). The results show that all treatment conditions (B, C, D) produced more problems and solutions than the control condition. 8 participants out of 12 participants produced more problems in the treatment conditions (B, C, D) than in the control condition (A). 10 participants out of 12 produced more solutions in the treatment conditions (B, C, D) than in the control condition (A). A common pattern in all conditions is that all participants produced more problems than solutions.

As shown Figure 17, the results of the group A&B did not show a distinct pattern between condition A and B. While P1, P7, P10 produced more problems and solutions in condition B than in condition A, P4 produced fewer problems and solutions in condition B than in condition A and P7 produced a similar number of problems and solutions in condition A and B. In addition, all participants produced much more problems than solutions.



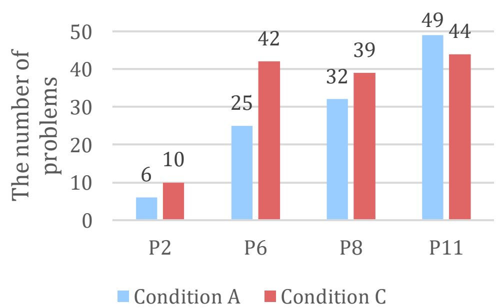
The number of problems (R+F+Be) in the group of A&B



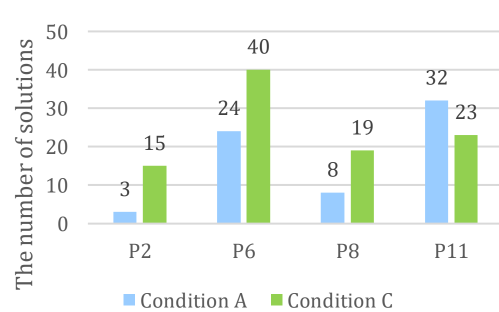
The number of solutions (Bs+S) in the group of A&B

Figure 17: The number of problems and solutions in the group of A&B

The results of the group A&C show that three out of four participants produced more problems and solutions in condition C than in condition A as shown in Figure 18. Especially, the results of the Solution show a large difference between the condition A and C.



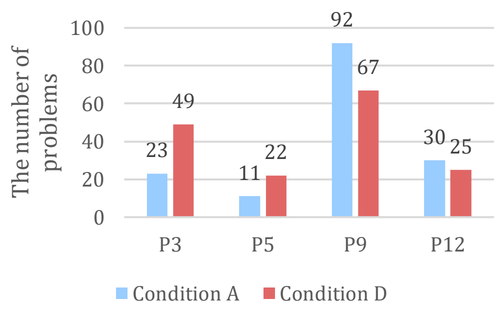
The number of problems (R+F+Be) in the group of A&C



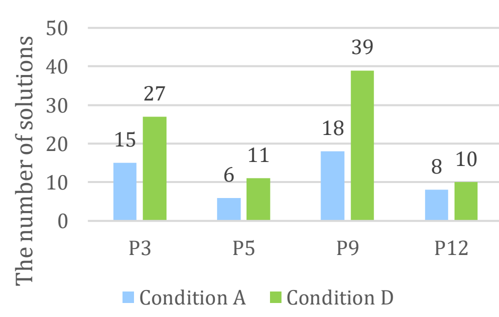
The number of solutions (Bs+S) in the group of A&C

Figure 18: The number of problems and solutions in the group of A&C

The results of the group A&D show that all participants produced more solutions in condition D than in condition A as shown in Figure 19, indicating the effect of sketch inspirations that are conceptually similar. For Problem, although P3 and P5 produced more problems in condition D than in condition A, P9 and P12 produced fewer problems in condition D than in condition A.



The number of problems (R+F+Be) in the group of A&D



The number of solutions (Bs+S) in the group of A&D

Figure 19: The number of problems and solutions in the group of A&D

4.3 Discussion and Research Implications

While most evaluations of existing co-creative systems focus on evaluating the usability [49], we focus on evaluating how a co-creative agent influences ideation in a human-AI collaboration. To evaluate the effect of AI inspiration on ideation, we applied an outcome-based approach (i.e. novelty, variety, quality, quantity) and a process-based approach (i.e. P-S index) in this pilot study. Overall our findings show that the AI-based stimuli produce different ideation outcomes and processes when compared to random stimuli. More specifically, we found that different types of AI-based stimuli show potential for different types of ideation. Novel ideation is associated with AI-based conceptually similar stimuli. Idea variety and quantity is associated with both AI-based visual and conceptual similarity of the inspiration. Idea quality is associated with visual similarity. 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.

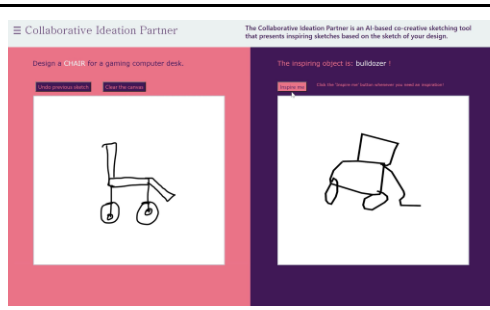
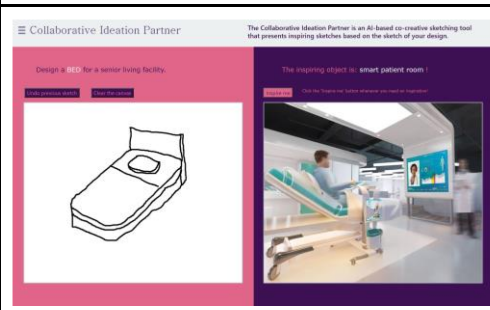
This pilot study does not have a sufficient number of participants to allow us to generalize the results for all cases of ideation from AI-based visual and conceptual similarity. However, we did a significance test on the results to see if there are significant trends to look for in a more robust study. A paired t-test was conducted to determine the significance of our results between the control condition and the treatment conditions in novelty, variety, and quantity. The results showed a significant difference in variety and quantity. For variety, participants in condition C ($M=24.25$, $SD=10.21$) produced more functions than in condition A ($M=15.50$, $SD=9.81$), $t(3)=-5.14$, two tail $p=0.014253$. For quantity, participants in condition B ($M=28.75$, $SD=12.76$) produced more functions than in condition A ($M=19.00$, $SD=13.24$), $t(3)=-3.30$, two tail $p=0.045732$. This pilot study does not have enough participants to measure or check for statistical significance, but the trends of the results show the potential for further analysis of the effect of an AI model for visual and conceptual similarity on design ideation with the metrics we identified for measuring ideation.

In our observations of the pilot 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.

CHAPTER 5: COLLABORATIVE IDEATION PARTNER (CIP DESIGN)

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. To improve the CIP system and 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. Table 7 shows the comparison between the CIP system used for the exploratory study (i.e. CIP Sketch) and the updated CIP system (i.e. CIP Design). The CIP Design focuses on conceptually similar inspirations to the target design and provides high fidelity images of creative designs.

Table 7: Comparison of CIP Sketch and CIP Design

	CIP Sketch	CIP Design
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	4 modes: random, similar, conceptually	2 modes: random inspiration,

inspiratio n	similar and visually different, visually similar and conceptually different	conceptually similar
Dataset	3450 Sketches (QuickDraw [47])	100 images

5.1 System Overview

The CIP Design is a co-creative sketching system which builds on the pilot study of the CIP Sketch that interprets sketches drawn by a user and provides inspirational sketches based on visual similarity and conceptual similarity. We developed the CIP Design to explore the effect of an AI Model for conceptual similarity on ideation during a design sketching task. The user interface of the CIP design is similar to the CIP Sketch as shown in Figure 20. The major difference between the CIP sketch and the CIP design is that the CIP design user interface provides an inspiring image of a creative design instead of a simple sketch. For example, in Figure 20, 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.

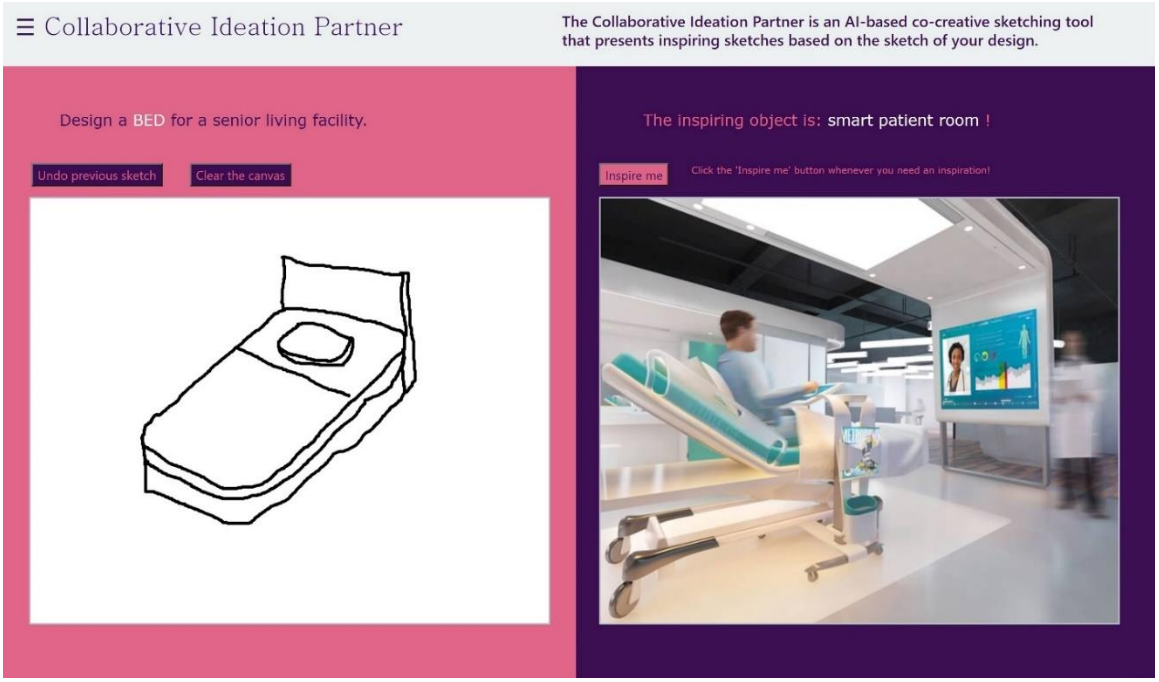


Figure 20: User Interface of the Collaborative ideation Partner Design

5.2 Design Image 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 [47] 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 “creative”, “novel”, “unusual”, “design” (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). These features were assigned by the research team.

5.3 AI Models 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 Wikipedia pre-trained word2vec model 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) 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.

5.4 User Experience: AI-Based Inspiration in CIP Design

Figure 21 shows an example of how a user communicates with the CIP system during design ideation. As shown in Figure 21, the participant drew a basic bed, a rectangular bed shape on wheels, without any special functions for the context of a senior living facility before requesting the first inspiration from the AI partner. The first inspiring image was a smart patient room and the participant responded by adding a reclining function with the foldable bed back. The participant then requested a second inspiring image which is a balcony pool. The participant added a foot pool as a new function, in response to the balcony pool image. The participant described that *"For the balcony pool, the first thing* *that I was thinking of was those multiple layers of pool. So I drew a shelf underneath the bed for some storage and a second layer, but then I was thinking it might also be nice to have a foot pool, you could sit up on the bed and put your feet in like the warm water. So then I added that foot pool to the end of the bed."*. The participant then requested the third inspiring image which is a wearable sleeping bag. From the wearable sleeping bag image, the participant came up with a built-in blanket idea and added it to the side of the bed. The participant described that *"I wasn't exactly sure how to use those sleeping bags, but I decided there could be like a built-in zippable cover maybe so I sort of drew on the side of the bed, like a bed roll, that's attached and could unfold and wrap over top of you. Built in blankets"*. The fourth inspiring image is a modular office table. From the modular office table image, the participant simply mimicked the table as a bedside table. In this case, the idea came from the inspiring image transferred to both the new function of the target design and the shape of the target design. The participant described that *"For this Modular Office table, I was just thinking that it'd be nice to have, like a table that could be a bedside table if you're laying down it could swivel in front of you if you're sitting up it'd be more of a tray table industry that in a similar shape to the Modular Office table the actual."*. From the fifth inspiring image which is an emergency tent, the participant came up with a canopy for privacy. The participant described the response to the emergency tent: *"And then looking at the emergency tent. I just thought it might be nice to have a little bit of privacy so you can roll your bed around the facility and also close yourself up so I added a little canopy to the top of the bed so that you could close the curtains if you wanted to."*. In summary, the interaction between the participant and the AI partner during ideation shows the participant mimicking a function or a shape from the inspiring images (i.e. 1st, 4th inspiring image) and introducing new functions by reinterpreting the concept of inspiring objects (i.e. 2nd, 3rd, 5th inspiring image).

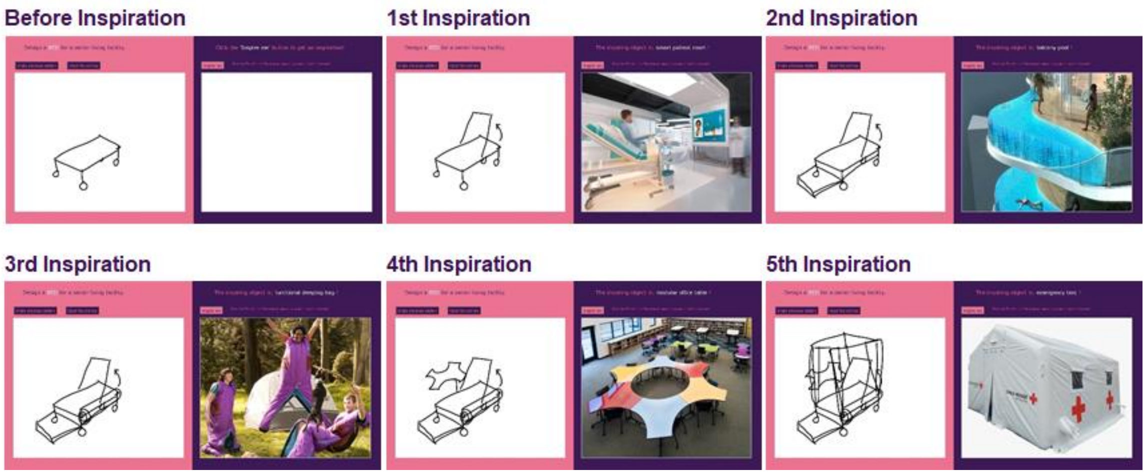


Figure 21: Example of Ideation Process Using CIP

CHAPTER 6: USER STUDY OF IDEATION WITH THE CIP DESIGN

This user study of the CIP Design focused on identifying distinct patterns of the participant's ideation in a human-AI collaboration where the AI partner contributes content based on conceptual similarity.

6.1 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), a treatment condition with conceptually similar inspirations.

- 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), 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. 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 images inspired their design in the retrospective protocol session.

6.2 Analysis of Data Collected

To measure ideation in a co-creative system, we developed three metrics based on [76], used for evaluating idea generation in design: novelty, variety, and quantity of design. We define the effect of the co-creative system as contributions of the AI agent to the idea generation. Two basic criteria are identified to define the contributions of the AI agent based on [76]:

- How well does the AI agent contribute to expanding the design space?
- How well does the AI agent contribute to exploring the design space?

We employ two approaches with the three metrics: an aggregated approach and a temporal approach. The aggregate approach allows us to evaluate the contributions of an AI agent in a design ideation. The representation of ideation process by temporal changes of ideas allows to (1) compare an ideation process of a design session to other design sessions, (2) identify specific patterns of novelty, variety, and quantity of ideas in a condition, (3) identify specific contributions of the co-creative system associated with novelty, variety, and/or quantity.

6.2.1 Data Collected

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 entire 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. From the recordings of retrospective sessions, we collected verbal data in which the participants explain what they were thinking and doing during the 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.

6.2.2 Data Segmentation and Coding

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 [32, 35]. We further segmented the inspiring segments until each segment has a single code in the FBS ontology [32, 35]. 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.

Table 8: Example of data segmentation and coding

Name of inspiring image	Inspiring image segment	Idea segment	Idea (stemming)
functional sleeping bag	I wasn't exactly sure how to use those sleeping bags, but I decided there could be like a built-in zippable cover maybe so I sort of drew on the side of the bed, like a bed roll, that's attached and could unfold and wrap over top of you. Built-in blankets, I guess.	I wasn't exactly sure how to use those sleeping bags, but I decided there could be like a built-in	built-in
		zippable	zip
		cover	cover
		maybe so I sort of drew on the side	side
		of the bed,	bed
		like a bed roll,	bed roll
		that’s attached	attach
		and could unfold	unfold
		and wrap over	warp
		top of you.	top
		Built-in	built-in
		blankets, I guess.	blanket

Table 8 shows an example of data segmentation and coding for an inspiring image. The example includes an inspiring image segment (i.e. 2nd column) for the “functional sleeping bag” image (i.e. 1st column) and the inspiring image segment is segmented into 12 idea segments (i.e. 3rd column). Each idea segment is then summarized as a single word that describes the idea (4th column). 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 then came to consensus for the different coding results. The inter-rater reliability (IRR) using the formula described in Miles and Huberman [64] was 0.98 which indicates sufficient agreement among the two coders.

6.2.3 Interpretation of Coded Data: Aggregated Analysis

To measure ideation in the design sessions, we developed three metrics based on the measurement of ideation in [76]: novelty, variety, and quantity of design. These metrics provide a basis for evaluating the effect of the AI inspiration on exploring (variety and quantity) and expanding (novelty) the design space [76].

The transcription data comprises 110 sessions of retrospective protocol (i.e. 55 sessions of condition A, 55 sessions of condition B). 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 9. This result indicates that AI inspiration based on conceptual similarity is more effective in design ideation than random inspiration. This section describes how we define and measure the contribution of the AI agent with the metrics of novelty, variety, and quantity.

Table 9: 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. In this study, 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 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.

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

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 condition A, as shown in Figure 22. 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 the contribution of the AI agent based on conceptual similarity in design ideation is more effective in novelty of ideas than when co-creating ideas with random inspirations.

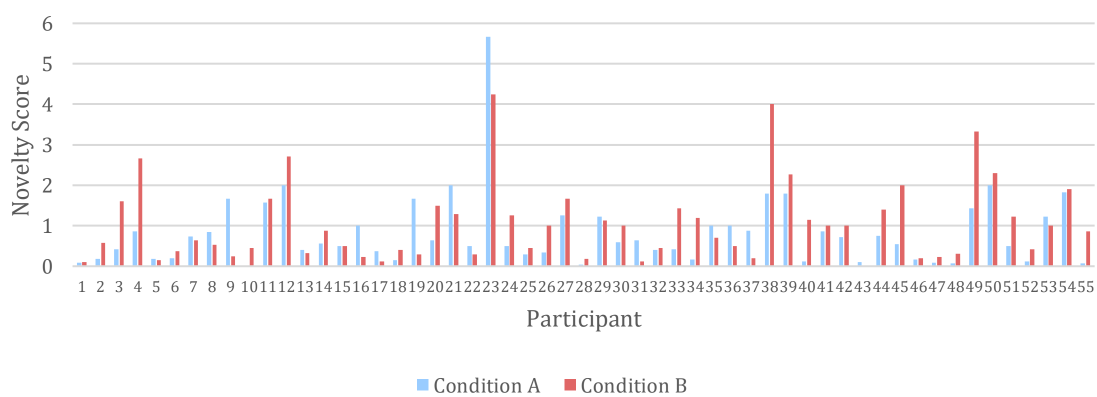


Figure 22: Novelty scores of participants

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. While the participants in condition A generated many similar ideas regardless of random inspirations, the participants in condition B showed generated ideas that were common to other participants in the same condition and unique ideas. For example, the most conceptually similar image in condition B is a smart patient room which includes an adjustable bed with wheels, foldable side table, bed rail guard, magazine holder, trash basket, and health information on a large display. The smart patient room image is not only conceptually similar to the target design object (i.e. bed) but also to the context of the design task (i.e. senior living facility). From the smart patient room image, many participants borrowed a function of the design (e.g. reclining function, movable bed, foldable large display) or mimicked a useful object directly (e.g. bed rail guard, bedside table, magazine holder, trash basket) for their design. In the retrospective sessions, the participants mostly described that the smart patient room is strongly associated with their design and helped a lot for developing their design. For example, P2 said *"The first image I got the smart patient room helped a lot."*; P4 said *"So in this next picture, I found a ton of inspiration from this picture, because the smart patient room it looks almost exactly like a hospital room, or like seniors who would be at for extended amounts of time."*; P12 said *"So as I clicked the first inspiring portion lucky enough it came up at a doctor's office or a patient's room. So that was a lot there I really need right there."*.

In addition to generating common ideas from the same inspiring image, participants in condition B often generated novel ideas, therefore expanding the concept of inspiring images. Figure 23 shows an example of a novel idea and how the participant thinks about the conceptually similar inspiring image to generate a novel idea. The inspiring image in Figure 23 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 expanding the concept of robotic arm. P50 transferred the robotic arm concept to another type of a robot with different purposes, a robotic bed which has robotic legs for mobility. P50 described that *"And then like the robotic arm. This bed is a little bit more like Boston Dynamics sort of those walking with, you know, big dog or spot. And so, that* *makes it to the bed will be able to access different environments. Easier than just a bed that has a track and two photo sensors or whatever. And so you would be able to get outside to those beach scenes in the other ones.”*. Once a robotic walking bed was created, P50 started to expand the function of the robotic 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. P50 described that *“And so, this is the robotic arm coming up I was thinking, that would be able to deploy the umbrella that, for whatever reason, or in this one, it's just a light. So thinking it could just be a multi tool sort of arm can move around the bed. That's the little umbrella thing. And it's just an attachment of that, that arm or bed tool.”*. This example indicates that the AI-based inspiration of CIP can support generating novel ideas in design ideation expanding the design space.

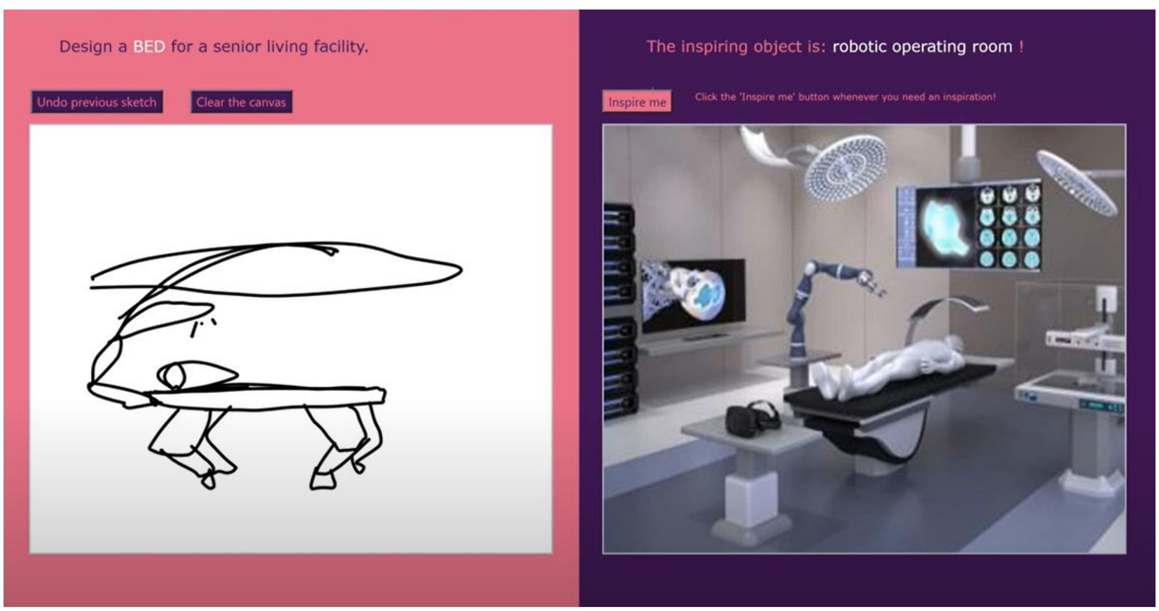


Figure 23: Example of novel idea that P50 generated

Variety. Variety is a measure of the explored solution space during the idea generation process [76]. For measuring variety, we code each idea whether it is a new idea or a repeated idea in a design session and only the number of new ideas is counted in a design session. A repeated idea is counted only one time as a new idea to identify the variety of ideas. 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 rest of three repeat ideas 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).

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

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 condition A, as shown in Figure 24. A paired t-test was conducted to determine the significance of the result between the control condition and the treatment condition in variety. A paired t-test was conducted to determine the significance of our results between the control condition and the treatment conditions 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 the contribution of the AI agent based on conceptual similarity in design ideation is more effective in producing a variety of ideas than when co-creating ideas with random inspirations.

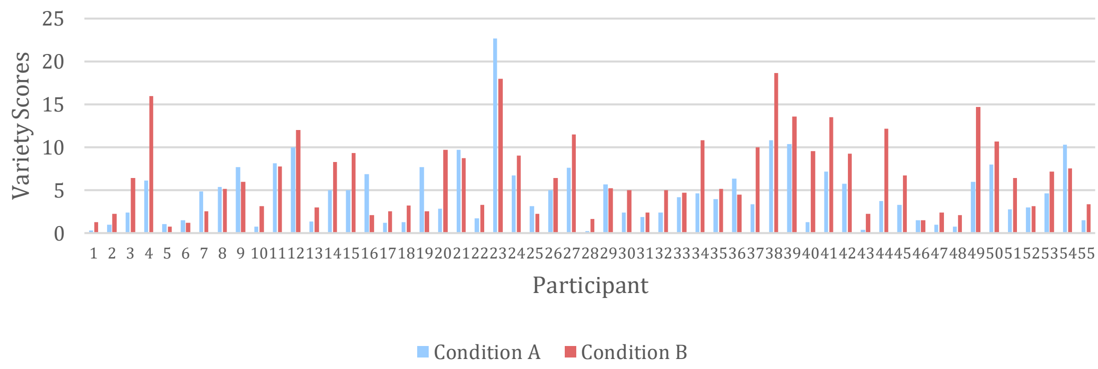
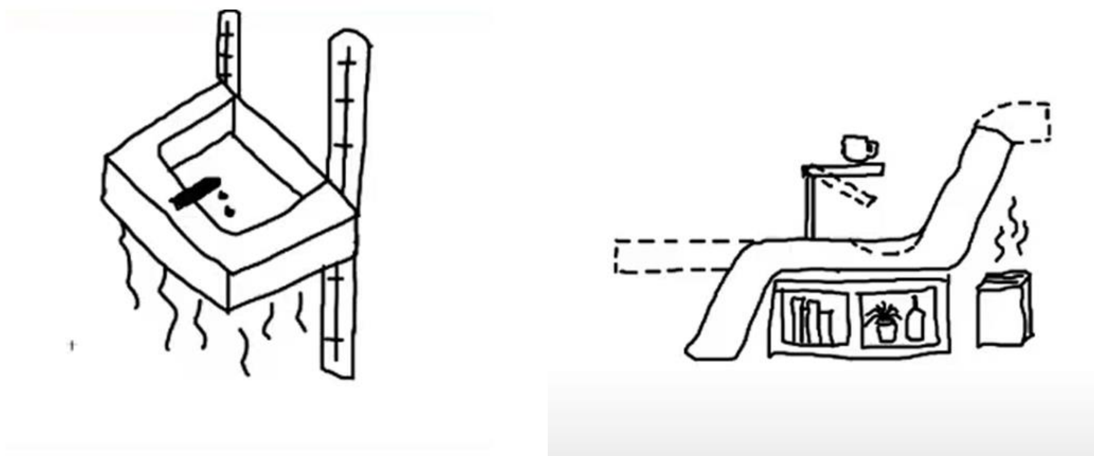


Figure 24: Variety scores of participants

From the observation of the video stream data, we found that participants tend to generate various ideas associated with the details of design and the context of the design getting inspirations from the CIP. Figure 25 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 and the context. To specific, the sink design in condition A includes three major functions (i.e. height adjustable sink, faucet on the user side, hand dryer under the sink) and the bed design in condition B includes four major functions (i.e. reclining bed, foldable side table, bookshelf under the bed, portable heater). Neither design has many functions, but the sketch of the sink design shows the main function concisely, while the sketch of the bed design shows the interaction with the design by expressing the folding angle through the dotted line (i.e. side table, leg part of the bed, seating part of the bed, head part of the bed) and the context of the design expressed through the surrounding objects (i.e. book, plant, coffee cup, vase).

P38 in condition B produced most of the new ideas when inspired by the smart patient room image, the most conceptually similar image to the target design. P38 came up with a reclining bed idea before seeing the smart patient room image but refined the initial idea with the inspiring image. P38 described that *"And then, Yeah, I got this first inspiring image, which pretty similar to what I drew. But I was also thinking if they're sitting in bed all day. Leaning their head back or maybe they're like, kind of tall or something, they might need some neck support. So that's what that dotted area is representing."*. P38 also noticed some useful objects (i.e. table, garbage bin, magazine stand) for additional functions of the bed in the inspiring image and this helped P38 to generate various ideas associated with new features of the bed and expressing the context of new features. P38 described that *"And then this image has. They have their like own table and own little bins for garbage I'm assuming and some little kind of stand for magazines, so I'm thinking about storage, and having space like for them to have something. So I made like a little cubicle underneath the bed and drew some random stuff in there, some books, plant, and a bottle, I think. Yeah. Last part of is that I think I still need something for them to be able to write on or like hold things on. Um, I couldn't think of a more modern way than to just have it come from the side like it does in this image. So after this I drew a little table. There's a plant and the bottle. Then now about to start drawing the last part which is that table. But I kind of see it as like a swivel table that can rotate at least in the, like, horizontal direction like the x axis, it can rotate if they get closer or further away for whatever reason."*. In this case, the table, garbage bin, magazine stand in the inspiring image were transferred to a storage idea in P38's design and the storage idea led to various objects to express the context of the storage. P38 also borrowed a table idea from the inspiring image but refined the idea considering the directions of the swivel table. This example shows that the AI-based inspiration of CIP can support generating a variety of ideas associated with the details of design and the context of the design.



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

Figure 25: Example of variety of ideas that P38 generated.

Quantity. In this study, the quantity of ideas is defined as the total number of times any idea that is activated in cognitive issues. For measuring quantity, the number of ideas both new ideas and repeated ideas is counted in a design session and is divided by the number of inspirations, as shown in Equation (3).

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

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 condition A, as shown in Figure 26. 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 the contribution of the AI agent based on conceptual similarity in design ideation is more effective in quantity of ideas than when co-creating ideas with random inspirations.

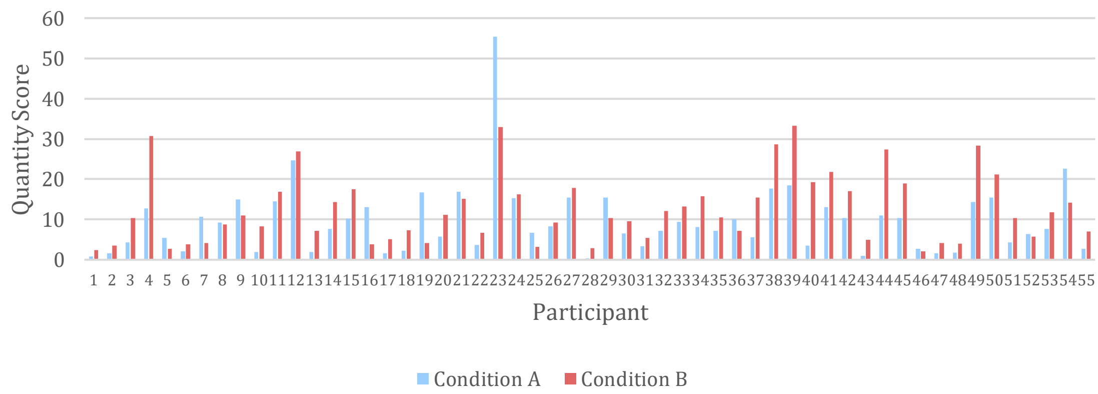
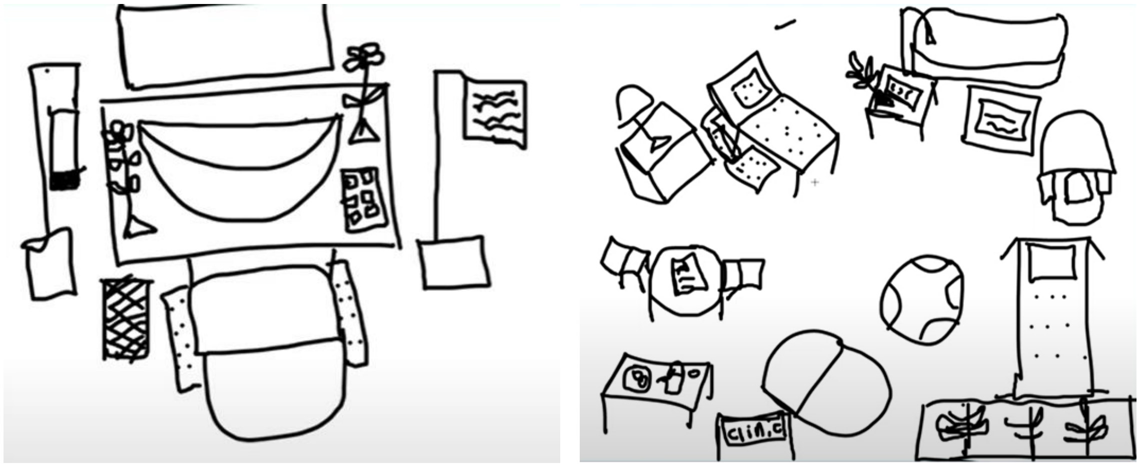


Figure 26: Quantity scores of participants

In examining the result of quantity scores, some participants showed a large difference between the control condition and the treatment condition. For example, P55 had a quantity score in condition A: 2.77 and a quantity score of condition B: 7.04. Figure 27, 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) Condition A (quantity score: 2.77) (b) Condition B (quantity score: 7.04)

Figure 27: Example of quantity of ideas that P55 generated.

A common design pattern of P55 in condition A and condition B is to include many individual items such as chair, table, remote control, and plant. However, the bed design of condition B shows a much larger number of items than the sink design of condition A. This is one of the major reasons that led to a large difference of idea quantity between the control condition and the treatment condition since many ideas were based on the individual items such as the purposes and the functions of the individual items. In condition B, P55 borrowed many items from inspiring images by directly copying items from an inspiring image or interpreting the features of an inspiring image. For example, P55 borrowed a remote control, armrest, and tv from a smart patient room image (i.e. *"So what I thought was, it would be good to have some sort of remote control with the bed that comes attached to the bed. So I'm drawing the armrest, that's also attached to a remote control. And then I drew. That was a TV, where they can watch anything they want. Inside the room just to relax."*); tables and chairs from a modular office table (i.e. *"And then when I saw this common area classroom thing (modular office table). I said why not have another chair, or another table, where I start drawing a table."*); and clinic place from an emergency tent image (i.e. *"and then I drew for some inspiration, because I wanted to add some more stuff. Then, I wanted to draw just some sort of a clinic, some sort of mini hospital for them to go to get their check up done stuff but I haven't enough space."*). This example indicates that the AI-based inspiration of CIP can support generating many ideas through facilitating the exploration of design spaces in design ideation.

In this section we explored the results of the participants' data in each of the categories of novelty, variety, and quantity of ideas for the 2 conditions. We show that in each of these categories the results of the treatment condition showed higher scores and a significant difference than the control condition. In considering combinations of the 3 categories, we found that 30 participants in the treatment condition produced higher scores in all novelty, variety, and quantity than in the control condition, while 10 participants in the control condition produced higher scores in all novelty, variety, and quantity than in the treatment condition; 12 participants in the treatment condition produced higher scores in two out of novelty, variety, and quantity than in the control condition; 3 participants in the treatment condition produced higher scores in one out of novelty, variety, and quantity than in the control condition. These findings show that there are different patterns and combinations of increasing novelty, variety, and quantity in the treatment condition, indicating that our hypothesis is confirmed.

6.2.4 Interpretation of Coded Data: Temporal Analysis

We employ a temporal analysis of ideation to test our hypothesis. The temporal analysis enables a characterization of the flow of ideas during a design session. We divide a design session as a series of segments bounded by the input of inspiration from the AI agent. For the temporal analysis, the number of novel ideas, the variety of ideas, and the quantity of ideas are calculated for each segment to produce a sequence of temporally ordered ideas in a design session. The nuances of the ideation process are then illustrated by temporal changes of novelty, variety, and quantity of ideas over time. The representation of ideation process by temporal changes of ideas allows to (1) compare an ideation process of a design session to other design sessions, (2) identify specific patterns of novelty, variety, and quantity of ideas in a condition, (3) identify specific contributions of the co-creative system associated with novelty, variety, and/or quantity.

The participants had 704 inspiring images in Condition A and 583 inspiring images in Condition B. The largest number of inspiring image segments among the 55 design sessions in Condition A is 50 segments (i.e. "before inspiring image segment" and 49 inspiring image segments). The largest number of inspiring image segments among the 55 design sessions in Condition B is 30 segments (i.e. "before inspiring image segment" and 29 inspiring image segments). For the temporal analysis, we thus calculated the number of novel ideas, the variety of ideas, and the quantity of ideas for each segment of 50 segments in Condition A and for each segment of 30 segments in Condition B. However, we compare the results of the first 30 segments between Condition A and Condition B in this analysis since (1) the results after 30 segments in condition A showed only a few ideas and (2) comparing the same number of segments shows a clear comparison of the temporal patterns of Condition A and Condition B.

We hypothesize that the quantity in novelty, variety, and quantity of ideas over time decreases more slowly in an ideation with AI-based inspirations based on conceptual similarity than the temporal pattern of ideation with random inspirations in a creative design task. Prior research shows a time effect on idea generation that the rate of idea generation decreases over time [42, 52, 69, 77, 80]. We thus posit that the effect of AI-based inspirations based on conceptual similarity can decrease the time effect on design ideation.

For the temporal analysis of novelty, we counted the number of novel ideas for each inspiring image segment. Figure 28 shows the number of novel ideas for the first 30 segments in Condition A and Condition B. The results of temporal patterns show that the quantity of novel ideas decreases over time in both Condition A and Condition B as we expected a time effect. However, there are two distinct patterns between the results of Condition A and Condition B. The changes of idea number between the "before" segment and the 1st segment shows a contradictory result between Condition A and Condition B. In condition A, the number of ideas was largely decreased in the first segment. However, in Condition B, the number of ideas was largely increased in the first segment. In the “before” segment, participants mostly produced many ideas without any inspiring images since they generate ideas to meet the basic requirements of the target design and to create a basic concept of their design. This result indicates that the conceptually similar image leads to more novel ideas. The second pattern we found is the distribution of the ideas. Although the results of both Condition A and Condition B show a decreasing pattern, the distribution until the 20th segment shows a significant difference between Condition A and Condition B.

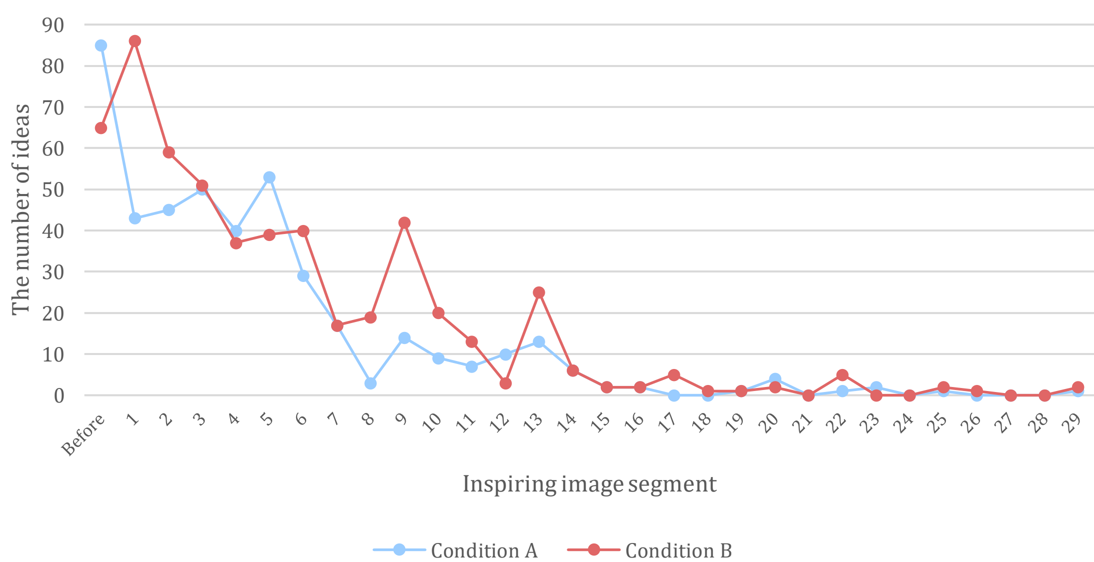


Figure 28: The number of novel ideas in each inspiring image segment.

To visualize the differences between the temporal patterns of Condition A and B, we generated trend lines with a scatter plot and the results showed a clear difference between

Condition A and Condition B as shown in Figure 29. A paired t-test was conducted to determine the significance of the result between the control condition and the treatment condition in temporal changes of novelty. We compared the results of trend line values between Condition A and Condition B and the results showed a significant difference between the control condition and the treatment condition. The quantity of novel ideas over time decreased more slowly in Condition B (M=18.17, SD=21.52) than in Condition A (M=14.60, SD=19.68), t(54)=1.69, two tail p=2.46E-11.

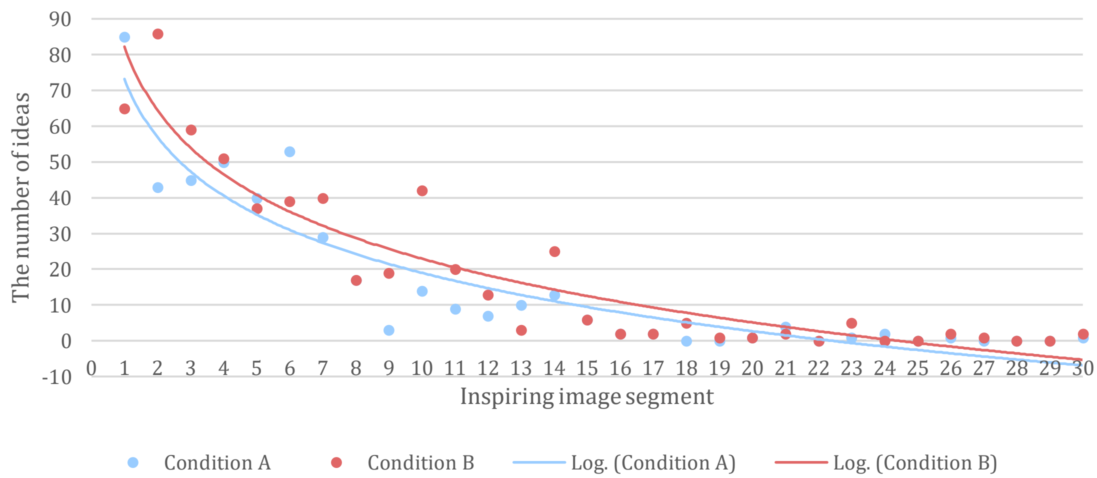


Figure 29: The trendlines of novelty.

For the temporal analysis of variety, we counted the number of new ideas for each inspiring image segment. Figure 30 shows the number of new ideas for the first 30 segments in Condition A and Condition B. The results of temporal patterns are similar to the results of novelty in terms of decreasing patterns with the time effect and the increasing pattern of Condition B in the first segment. However, the increasing pattern of Condition B in the first segment is relatively less than the results of novelty and the result of Condition B after the first segment shows a similar pattern to Condition A largely decreasing in the second segment.

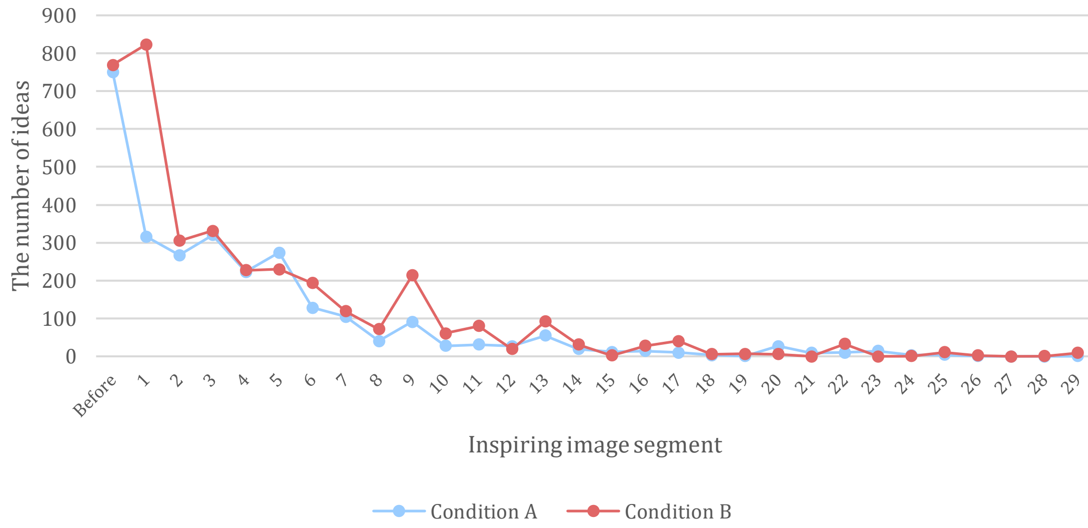


Figure 30: The number of new ideas in each inspiring image segment.

Figure 31 shows the results as trend lines with a scatter plot for the temporal patterns of variety. The results mainly show a difference from the first segment to the 20th segment and do not show a large difference after the 20th segment. A paired t-test was conducted to determine the significance of the result between the control condition and the treatment condition in temporal changes of variety. We compared the results of trend line values between Condition A and Condition B and the results showed a significant difference between the control condition and the treatment condition. The quantity in variety of ideas over time decreased more slowly in Condition B (M=123.78, SD=185.79) than in Condition A (M=92.79, SD=144.15), t(54)=1.69, two tail p=0.00039.

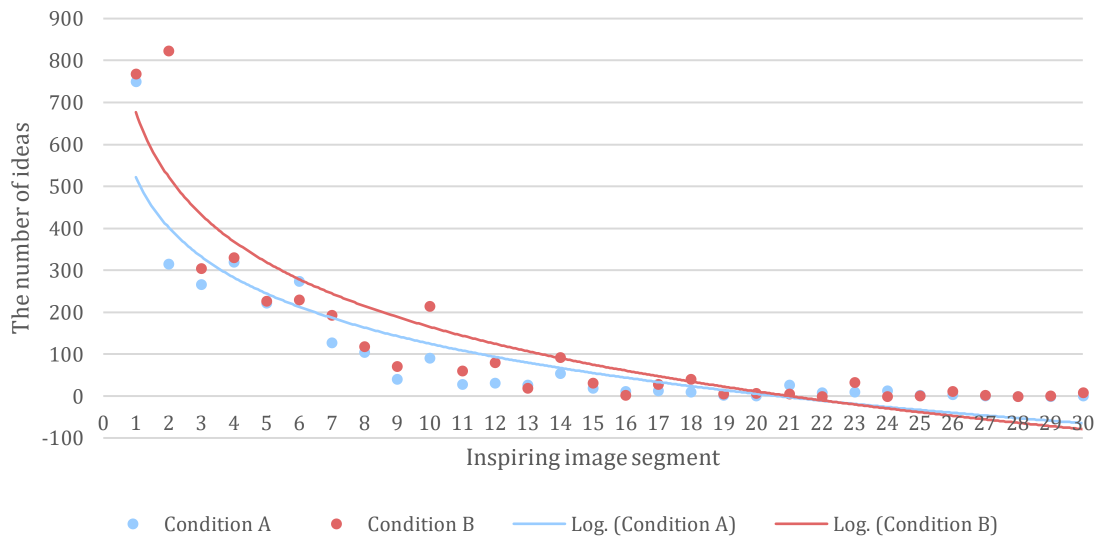


Figure 31: The trendlines of variety.

For the temporal analysis of quantity, we counted the total number of ideas including new ideas and repeated ideas for each inspiring image segment. Figure 32 shows the number of ideas for the first 30 segments in Condition A and Condition B. The results of temporal patterns are similar to the results of novelty and variety. In Condition B, the number of ideas was largely increased in the first segment like the result of novelty. The results after the first segment show a similar pattern to the results of variety but show bigger differences between Condition A and Condition B.

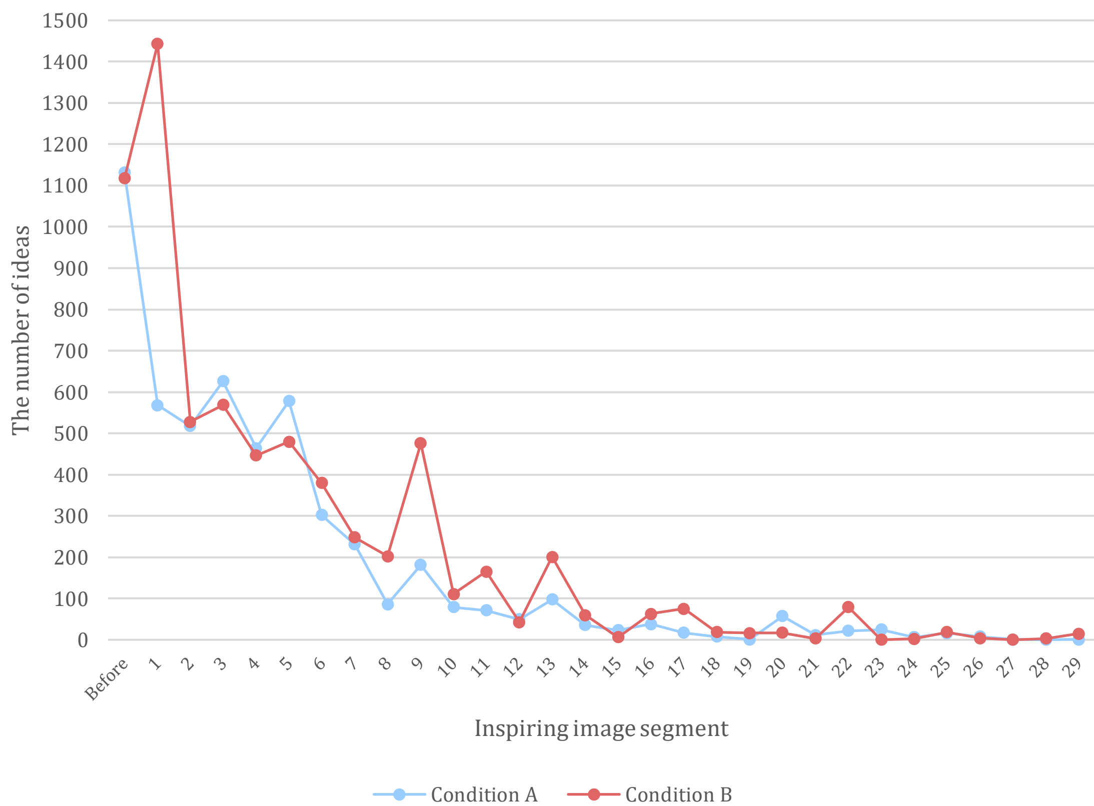


Figure 32: The number of ideas in each inspiring image segment.

Figure 33 shows the results as trend lines with a scatter plot for the temporal patterns of quantity. The results show a similar pattern to the results of variety, showing a difference from the first segment to the 20th segment. The result of a paired t-test that compares the trend lines showed a significant difference as well between the control condition and the treatment condition. The quantity in quantity of ideas over time decreased more slowly in Condition B (M=226.19.78, SD=307.37) than in Condition A (M=175.24, SD=248.33), t(54)=1.69, two tail p=0.000067.

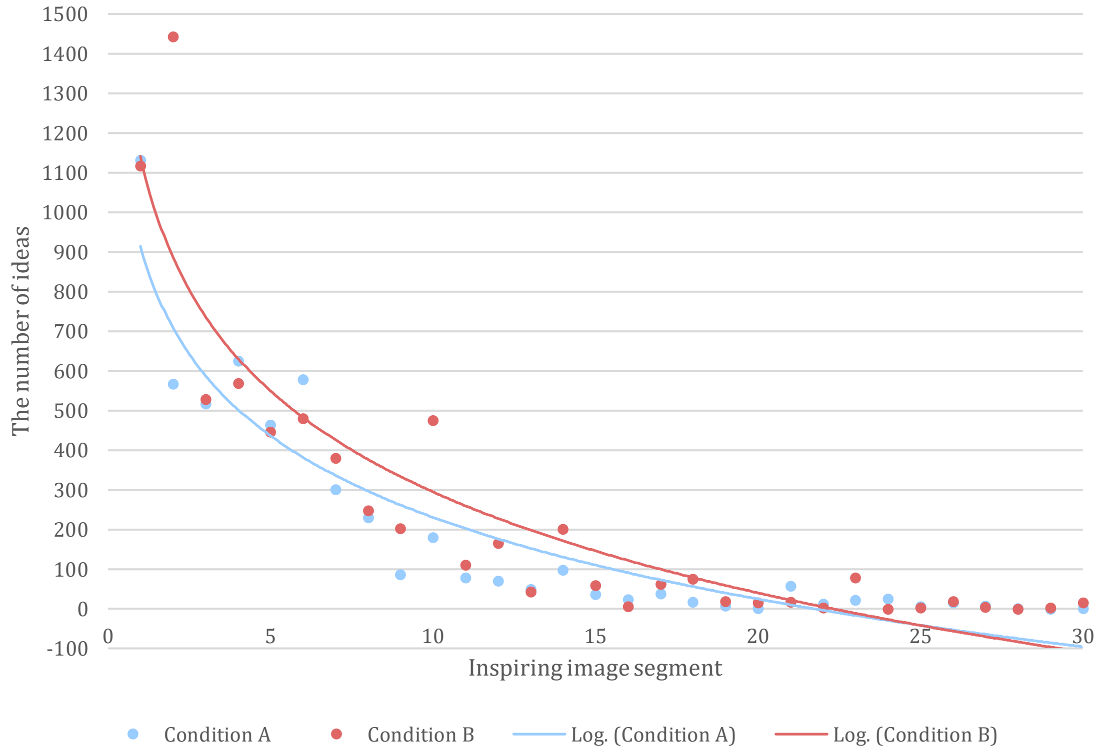


Figure 33: The trendlines of quantity.

CHAPTER 7: CONCLUSION AND FUTURE WORK

The aim of this thesis is to identify the effect of an AI model for conceptual similarity on design ideation in a co-creative sketching tool. Co-creative systems can support design creativity by encouraging the exploration of design solutions in design ideation. This thesis posits that conceptual similarity can be a basis for provoking design ideation in Human-AI collaboration. This thesis presented a co-creative sketching tool called Collaborative Ideation Partner (CIP) that supports the idea generation of new designs by providing conceptual stimuli. It was presumed that the similarity of the AI partner's contribution will influence the design ideation process.

CIP Sketch was developed to support a pilot study that evaluates the effect of an AI model for visual and conceptual similarity on design ideation in a co-creative design tool. CIP Sketch is a co-creative sketching tool that provides inspirational sketches based on the visual and conceptual similarity to sketches drawn by a designer. To generate an inspiring sketch, the AI model of CIP Sketch computes the visual similarity 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. CIP Sketch used four different categories of ideation stimuli based on visual and conceptual similarities (i.e. random, visually and conceptually similar, conceptually similar, visually similar) to explore the effect of an AI model for visual and conceptual similarity on ideation. To evaluate the ideation on the use of CIP Sketch, the pilot study used both the outcome-based approach and process-based approach. From the outcome-based approach, the findings show that the AI model for visual and conceptual similarity is more effective than random inspirations in design ideation; novelty may be associated with the conceptual similarity of the AI model; variety and quantity may be associated with both visual and conceptual similarity of the AI model; quality may be associated with visual similarity; requirement and structure may be associated with the conceptual similarity of the AI model in variety and quantity; function and structure may be associated with the visual similarity of the AI model in variety and quantity. From the process-based approach, the findings show that both visual and conceptual similarity of the AI model is associated with a problem-focused designing style but produces more solutions than random inspirations. The findings suggest that conceptually similar inspiration that does not have strong visual similarity leads to more novelty, variety, and quantity during ideation. The findings also suggest that visually similar inspiration that does not have strong conceptual similarity leads to more quality ideas during ideation. Future AI-based co-creativity can be more intentional by contributing inspiration to improve novelty and quality, the basic characteristics of creativity.

Based on what I learned from the pilot study, CIP Sketch was changed to CIP Design that provides inspirational images based on conceptual similarity. To study the impact of CIP Design on design ideation, we performed a user study with two conditions for the AI inspiration: random inspiration and conceptual similar inspiration. To evaluate the effect of AI inspiration on ideation, I measured the ideas generated by the user with three metrics: novelty, variety, and quantity. The user study for measuring the effect of CIP design is designed to validate the following hypothesis:

- Hypothesis 1: *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.*
- Hypothesis 2: *The quantity in novelty, variety, and quantity of ideas over time decreases more slowly in ideation with AI-based inspirations based on*

conceptual similarity than the temporal pattern of ideation with random inspirations in a creative design task.

Hypothesis 1 was validated through an aggregate analysis showing statistically significant differences between the control condition and the treatment condition in the novelty, variety, and quantity scores. In addition to the statistical analysis of the individual novelty, variety, and quantity scores, I found patterns across the three metrics in the different conditions. I found that the participants tend to request more inspiring images in the control condition but produce more ideas in the treatment condition. From our observations of the retrospective sessions, I believe that the participants could not find the relationship between the design task and the random inspirations in the control condition and therefore did not continue the ideation process before requesting another inspiring image. To better understand the increases in the novelty, variety, and quantity of ideas in the 2 conditions, we reviewed the think aloud data from the retrospective sessions to identify the context in which these metrics increased. Hypothesis 2 was validated through a temporal analysis showing statistically significant differences between the control condition and the treatment condition in the novelty, variety, and quantity. From the temporal analysis, I found that the effect of conceptually similar inspirations slows the time effect on idea generation that the rate of idea generation decreases over time. The participants tended to request more inspiring images in the control condition but the ideas decreased more quickly than in the treatment condition. To be specific, the control condition and treatment condition showed a significant difference at the beginning of the ideation process, the difference appeared until 20 segments of inspiring images.

Measuring ideation when co-creating with an AI-based co-creative design tool enables the comparison and evaluation of the impact of different AI models on the user's cognitive process and the creative outcome. While many ideation measures focus on product-based approaches, this thesis focuses on a cognitive approach to better understand how a co-creative agent influences ideation in a human-AI collaboration. In order to measure ideation, we developed an approach for measuring ideation that has two components: an aggregate analysis and a temporal analysis. The aggregate analysis adapts existing quantitative metrics for ideation: novelty, variety, and quantity of ideas expressed in the design session. The temporal analysis shows the temporal changes of novelty, variety, and quantity of ideas based on the AI contributions. We applied these measures to evaluate the effect of an AI model for conceptual similarity on design ideation in a co-creative design system. These measures can be used in evaluating the impact of AI contributions in other co-creative systems that support design creativity.

This thesis provides a basis for further exploration of the impact of AI-based inspiration on design ideation. In this thesis, the AI-based inspiration condition presented images that were ordered based on the conceptual similarity of the description of the design task and the description of the images in the data set. This order from most similar to less similar shows that the rate of ideation, measuring novelty, variety, and quantity decreased more slowly than the control condition. There are two research directions that I am planning for future studies, which are as follows: improving the Collaborative Ideation Partner and further evaluations of the Collaborative Ideation Partner.

To improve the Collaborative Ideation Partner system based on what we learned from this thesis, I plan to measure conceptual similarity based on an updated description of the design rather than holding the design description constant as the task description. This allows the measure of conceptual similarity to be connected to the latest version of the design description. An idea for updating the design description is to create a list of design descriptions including design requirements and specific problems the designer should solve and update each design description sequentially. When the design description is updated, the measure of conceptual similarity is connected to the new design description and inspiring images are provided based on the updated design description. In this way, the designer can focus on a specific design issue with a design description and the researchers can observe how the designer solves the specific problem by communicating with AI-based inspirations.

Next, I plan to explore ways to extend the dataset to searching the web for images with attribution. This allows the AI to select from a range of similar images so it is not always presenting the most similar unseen inspiration. The foundational results of the user study show that AI-based inspiration based on conceptual similarity, even in this limited AI ability to measure similarity, improves ideation. The approach to measuring ideation allows us to evaluate the effect of additional considerations in the selection of the next inspiring image.

Last, I plan to improve the algorithm for visual similarity to address the increasing complexity problem. From the pilot study, we found some issues associated with AI-based visually similar inspirations. Although the results of the pilot study showed no significant differences between the control condition and the visually similar condition, we could not fully explore the effect of AI-inspirations based on visual similarity due to the limited AI ability of the CIP sketch. Based on the results of the pilot study and what we learned from the pilot study, we focused on conceptual similarity to developing the CIP design. However, this does not mean visual similarity is not important and relevant in ideation. To solve the complexity problem and apply AI-based visually similar inspirations to the CIP design, the CIP system needs a larger image dataset and an algorithm for Sketch-to-Image matching. In a future study, I plan to measure the effect of AI-based inspirations based on visual similarity with an improved algorithm for visual similarity.

Another research direction for future studies is further evaluations of the Collaborative Ideation Partner. This thesis compared AI-based conceptually similar inspirations to random inspirations to validate the effect of AI-based inspirations. In a future study, I plan to compare AI-based similarity to a human search of similarity in a google search. In the early design process, human search is a common design method to inspire ideas from existing designs such as precedent studies and case studies. The future study will be a between-subject design that compares AI-based similarity to a human search of similarity in a google search. The same design task, an open-end design task, will be used for each condition. The participants will be asked to perform the design task for 15 minutes which is a time limit to complete the task. Once the participants finish the design task, the participants are asked to explain what they were thinking based on watching their design session recording in a retrospective protocol session. From this retrospective protocol analysis, researchers can identify the characteristics of human search (e.g. what keywords the human designers search, what images human designers see, where the human designers inspire from) that can be applied to co-creative systems. The researchers can also compare the ideation effectiveness between the human search and AI-based similarity using the approach presented in this dissertation. Comparing AI-based similarity to a human search of similarity thus allows us to better understand the effect of AI-based inspirations in a practical way.

Another further study of the Collaborative Ideation Partner is evaluating the effect of interaction designs on Human-AI collaboration in design ideation. Interaction design in Human-AI collaboration plays an important role to provide better communication and collaborative experience between the user and AI agent. In the user study of Collaborative Ideation Partner, some participants did not actively collaborate with the AI partner focusing on their own ideas or forgetting to collaborate with the AI partner. The participants also thought of the AI partner as a support tool rather than a creative partner. This issue shows the importance of interaction design associated with users' collaborative experience, engagement, and immersion in Human-AI collaboration. Some potential ideas to improve current CIP interaction design are: providing messages to encourage collaboration with the AI partner (e.g. providing a message "I can help you, click "Inspire me" when the participant did not click the "Inspire me" button for 3 minutes); providing inspiring images after automatically detecting changes in the user's sketch; providing a set of inspiring images when clicking the "Inspire me" button. These ideas provide different collaborative experiences for Human-AI collaboration enhancing user engagement. This thesis compared different AI abilities with the same interaction design to validate the effect of AI ability. I plan to study the effect of interaction design by comparing different interaction designs with the same AI ability in the future. This approach allows us to evaluate the additional effect of co-creative systems in ideation and can be a basis for designing co-creative systems.

REFERENCES

[1] Akin, Ö. 1990. Necessary conditions for design expertise and creativity. *Design Studies*. 11, 2 (1990), 107–113.

[2] Amabile, T.M. 1982. Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*. 43, 5 (1982), 997.

[3] Atman, C.J., Chimka, J.R., Bursic, K.M. and Nachtmann, H.L. 1999. A comparison of freshman and senior engineering design processes. *Design studies*. 20, 2 (1999), 131–152.

[4] Biles, J. 1994. GenJam: A genetic algorithm for generating jazz solos. *ICMC* (1994), 131–137.

[5] Biles, J.A. 1998. Interactive GenJam: Integrating real-time performance with a genetic algorithm. *ICMC* (1998).

[6] Brooke, J. 1996. Sus: a "quick and dirty'usability. *Usability evaluation in industry*. 189, (1996).

[7] Brophy, D.R. 2001. Comparing the attributes, activities, and performance of divergent, convergent, and combination thinkers. *Creativity research journal*. 13, 3–4 (2001), 439–455.

[8] Chan, J., Siangliulue, P., Qori McDonald, D., Liu, R., Moradinezhad, R., Aman, S., Solovey, E.T., Gajos, K.Z. and Dow, S.P. 2017. Semantically far inspirations considered harmful? accounting for cognitive states in collaborative ideation. *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition* (2017), 93–105.

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

[10] Cherry, E. and Latulipe, C. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)*. 21, 4 (2014), 1–25.

[11] Chiu, I. and Shu, L.H. 2012. Investigating effects of oppositely related semantic stimuli on design concept creativity. *Journal of Engineering Design*. 23, 4 (2012), 271–296.

[12] Christensen, B.T. and Schunn, C.D. 2007. The relationship of analogical distance to analogical function and preinventive structure: The case of engineering design. *Memory & cognition*. 35, 1 (2007), 29–38.

[13] Colton, S., Goodwin, J. and Veale, T. 2012. Full-FACE Poetry Generation. *ICCC* (2012), 95–102.

[14] Colton, S., Halskov, J., Ventura, D., Gouldstone, I., Cook, M. and Ferrer, B.P. 2015. The Painting Fool Sees! New Projects with the Automated Painter. *ICCC* (2015), 189–196.

[15] Compton, K. and Mateas, M. 2015. Casual Creators. *ICCC* (2015), 228–235.

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

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

[18] Daly, S.R., Christian, J.L., Yilmaz, S., Seifert, C.M. and Gonzalez, R. 2011. Teaching design ideation. (2011).

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

[20] Davis, N., Do, E.Y.-L., Gupta, P. and Gupta, S. 2011. Computing harmony with PerLogicArt: perceptual logic inspired collaborative art. *Proceedings of the 8th ACM conference on Creativity and cognition* (2011), 185–194.

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

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

[23] Davis, N.M. 2017. *Creative sense-making: A cognitive framework for quantifying interaction dynamics in co-creation*. Georgia Institute of Technology.

[24] Davis, N.M. 2013. Human-computer co-creativity: Blending human and computational creativity. *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference* (2013).

[25] Dorst, K. 2019. Co-evolution and emergence in design. *Design Studies*. 65, (2019), 60–77.

[26] Dorst, K. and Cross, N. 2001. Creativity in the design process: co-evolution of problem–solution. *Design studies*. 22, 5 (2001), 425–437.

[27] Dorta, T. 2008. Design flow and ideation. *International Journal of Architectural Computing*. 6, 3 (2008), 299–316.

[28] Douglas, D.H. and Peucker, T.K. 1973. 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*. 10, 2 (1973), 112–122.

[29] Elgammal, A., Liu, B., Elhoseiny, M. and Mazzone, M. 2017. Can: Creative adversarial networks, generating" art" by learning about styles and deviating from style norms. *arXiv preprint arXiv:1706.07068*. (2017).

[30] Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C. and Wood, K. 2013. The meaning of “near” and “far”: the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design*. 135, 2 (2013).

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

[32] Gero, J.S. 1990. Design prototypes: a knowledge representation schema for design. *AI magazine*. 11, 4 (1990), 26–26.

[33] Gero, J.S. 2010. Generalizing design cognition research. *DTRS*. 8, (2010), 187–198.

[34] Gero, J.S., Jiang, H. and Williams, C.B. 2013. Design cognition differences when using unstructured, partially structured, and structured concept generation creativity techniques. *International Journal of Design Creativity and Innovation*. 1, 4 (2013), 196–214.

[35] Gero, J.S. and Kannengiesser, U. 2004. The situated function–behaviour–structure framework. *Design studies*. 25, 4 (2004), 373–391.

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

[37] Goel, A.K. 1997. Design, analogy, and creativity. *IEEE expert*. 12, 3 (1997), 62–70.

[38] Goel, V. and Pirolli, P. 1992. The structure of design problem spaces. *Cognitive science*. 16, 3 (1992), 395–429.

[39] Grace, K., Maher, M.L., Fisher, D. and Brady, K. 2015. Data-intensive evaluation of design creativity using novelty, value, and surprise. *International Journal of Design Creativity and Innovation*. 3, 3–4 (2015), 125–147.

[40] Hernandez, N.V., Shah, J.J. and Smith, S.M. 2010. Understanding design ideation mechanisms through multilevel aligned empirical studies. *Design Studies*. 31, 4 (2010), 382–410.

[41] Hoffman, G. and Weinberg, G. 2010. Gesture-based human-robot jazz improvisation. *2010 IEEE International Conference on Robotics and Automation* (2010), 582–587.

[42] Howard-Jones, P.A. and Murray, S. 2003. Ideational productivity, focus of attention, and context. *Creativity research journal*. 15, 2–3 (2003), 153–166.

[43] Jacob, M., Coisne, G., Gupta, A., Sysoev, I., Verma, G.G. and Magerko, B. 2013. Viewpoints ai. *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference* (2013).

[44] Jacob, M., Zook, A. and Magerko, B. 2013. Viewpoints AI: Procedurally Representing and Reasoning about Gestures. *DiGRA conference* (2013).

[45] Jia, L., Becattini, N., Cascini, G. and Tan, R. 2020. Testing ideation performance on a large set of designers: effects of analogical distance. *International Journal of Design Creativity and Innovation*. 8, 1 (2020), 31–45.

[46] Jiang, H., Gero, J.S. and Yen, C. 2014. Exploring designing styles using a problem–solution index. *Proc. Design Computing and Cognition '12 (Gero, JS, Ed* (2014).

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

[48] Kantosalo, A. and Riihiaho, S. 2019. Quantifying co-creative writing experiences. *Digital Creativity*. 30, 1 (2019), 23–38.

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

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

[51] Kennedy, E.B., Miller, D.J. and Niewiarowski, P.H. 2018. Industrial and biological analogies used creatively by business professionals. *Creativity Research Journal*. 30, 1 (2018), 54–66.

[52] Liikkanen, L.A., Björklund, T.A., Hämäläinen, M.M. and Koskinen, M.P. 2009. Time constraints in design idea generation. *DS 58-9: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 9, Human Behavior in Design, Palo Alto, CA, USA, 24.-27.08. 2009* (2009), 81–90.

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

[54] Liu, Y.-C., Chakrabarti, A. and Bligh, T. 2003. Towards an ‘ideal’approach for concept generation. *Design studies*. 24, 4 (2003), 341–355.

[55] Lubart, T. 2005. How can computers be partners in the creative process: classification and commentary on the special issue. *International Journal of Human-Computer Studies*. 63, 4–5 (2005), 365–369.

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

[57] Lund, A.M. 2001. Measuring usability with the use questionnaire12. *Usability interface*. 8, 2 (2001), 3–6.

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

[59] Maher, M. and Tang, H.-H. 2003. Co-evolution as a computational and cognitive model of design. *Research in Engineering design*. 14, 1 (2003), 47–64.

[60] Maher, M.L., Brady, K. and Fisher, D.H. 2013. Computational models of surprise in evaluating creative design. *Proceedings of the fourth international conference on computational creativity* (2013).

[61] Maher, M.L., Poon, J. and Boulanger, S. 1996. Formalising design exploration as co-evolution. *Advances in formal design methods for CAD*. Springer. 3–30.

[62] Mamykina, L., Candy, L. and Edmonds, E. 2002. Collaborative creativity. *Communications of the ACM*. 45, 10 (2002), 96–99.

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

[64] Miles, M.B. and Huberman, A.M. 1994. *Qualitative data analysis: An expanded sourcebook*. sage.

[65] Moreno, D.P., Hernandez, A.A., Yang, M.C., Otto, K.N., Hölttä-Otto, K., Linsey, J.S., Wood, K.L. and Linden, A. 2014. Fundamental studies in Design-by-Analogy: A focus on domain-knowledge experts and applications to transactional design problems. *Design Studies*. 35, 3 (2014), 232–272.

[66] Nelson, B.A., Wilson, J.O., Rosen, D. and Yen, J. 2009. Refined metrics for measuring ideation effectiveness. *Design Studies*. 30, 6 (2009), 737–743.

[67] Newell, A. and Simon, H.A. 1972. *Human problem solving*. Prentice-Hall Englewood Cliffs, NJ.

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

[69] Perttula, M.K. and Liikkanen, L.A. 2006. Structural tendencies and exposure effects in design idea generation. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (2006), 199–210.

[70] Poon, J. and Maher, M.L. 1997. Co-evolution and emergence in design. *Artificial Intelligence in Engineering*. 11, 3 (1997), 319–327.

[71] Ramer, U. 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer graphics and image processing*. 1, 3 (1972), 244–256.

[72] Rashel, F. and Manurung, R. 2014. Pemuisi: a constraint satisfaction-based generator of topical Indonesian poetry. *ICCC* (2014), 82–90.

[73] Rehurek, R. and Sojka, P. 2010. Software framework for topic modelling with large corpora. *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (2010).

[74] Reinig, B.A., Briggs, R.O. and Nunamaker, J.F. 2007. On the measurement of ideation quality. *Journal of Management Information Systems*. 23, 4 (2007), 143–161.

[75] Sarkar, P. and Chakrabarti, A. 2011. Assessing design creativity. *Design studies*. 32, 4 (2011), 348–383.

[76] Shah, J.J., Smith, S.M. and Vargas-Hernandez, N. 2003. Metrics for measuring ideation effectiveness. *Design studies*. 24, 2 (2003), 111–134.

[77] Snyder, A., Mitchell, J., Ellwood, S., Yates, A. and Pallier, G. 2004. Nonconscious idea generation. *Psychological reports*. 94, 3_suppl (2004), 1325–1330.

[78] Srinivasan, V., Song, B., Luo, J., Subburaj, K., Elara, M.R., Blessing, L. and Wood, K. 2018. Does analogical distance affect performance of ideation? *Journal of Mechanical Design*. 140, 7 (2018).

[79] Taylor, D.W., Berry, P.C. and Block, C.H. 1958. Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative science quarterly*. (1958), 23–47.

[80] Tsenn, J., Atilola, O., McAdams, D.A. and Linsey, J.S. 2014. The effects of time and incubation on design concept generation. *Design Studies*. 35, 5 (2014), 500–526.

[81] Veale, T. 2014. Coming good and breaking bad: Generating transformative character arcs for use in compelling stories. *Proceedings of the 5th international conference on computational creativity* (2014).

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

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

[84] Wiggins, G.A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*. 19, 7 (2006), 449–458.

[85] Wiggins, G.A. 2006. Searching for computational creativity. *New Generation Computing*. 24, 3 (2006), 209–222.

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