


Enhancing design concept diversity: multi-persona prompting strategies for large language models

Wangchuan Bradley Feng¹, Sébastien Hélié² and Jitesh H. Panchal ¹

¹*School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA*

²*Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA*

Abstract

Large language models (LLMs) are increasingly used to address real-world design problems, especially during the design ideation phase. Although LLMs hold substantial promise for concept generation, the understanding of how they can effectively assist designers in enhancing the diversity of design concepts is still limited. In this study, we set up different strategies for prompting multiple professional personas to the LLM for design concept generation, including (1) multiple prompts for concept generation in parallel, each with a professional persona, (2) a single prompt for concept generation with multiple professional personas, and (3) a sequence of prompts for concept generation and update, each with a professional persona. We formulate and test several hypotheses on the effectiveness of different strategies. All hypotheses are tested by constructing professional knowledge bases, selecting design problems and personas, and designing the prompts. The results suggest that LLMs can facilitate the design ideation process and provide more diverse design concepts when they are given multiple prompts in parallel, each with a professional persona, or given a sequence of prompts with multiple professional personas to generate and update design concepts gradually.

Keywords: conceptual design, idea generation, large language models, prompt engineering, design ideation

Received 28 February 2025

Revised 06 October 2025

Accepted 07 October 2025

Corresponding author

Jitesh H. Panchal

panchal@purdue.edu

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Des. Sci., vol. 11, e52

journals.cambridge.org/dsj

DOI: [10.1017/dsj.2025.10037](https://doi.org/10.1017/dsj.2025.10037)



1. Introduction

Design ideation is a creative stage in the design process where designers gather with open minds to produce various solution concepts to address design problems. This process aims for designers to extensively explore the design space, which is the space of all possible design concepts (Daly *et al.* 2016). The effectiveness of this process relies significantly on the knowledge, experience and creativity of human designers. However, these same factors can also limit their ability to thoroughly explore the design space. Various methods have been developed to aid designers in generating concepts during the ideation phase (Osborn 1957; Daly *et al.* 2011; Álvarez & Ritchey 2015).

As large language models (LLMs) such as ChatGPT have evolved, several researchers have explored their potential to enhance the creativity of human designers (Filippi 2023; Girotra *et al.* 2023; Wang *et al.* 2023; Zhu & Luo 2023). These studies primarily focus on how LLMs can improve the efficiency of generating design concepts, while their impact on the diversity of generated concepts

remains underexplored. We have previously conducted preliminary research on how ChatGPT can be prompted with a professional persona to increase the diversity of generated design concepts (Feng, H  lie & Panchal 2024). Building on that work, this study addresses the following research question:

How can LLMs be prompted to maximize the diversity of design concepts generated during the design ideation process?

To answer this question, we formulate a central hypothesis and sub-hypotheses:

Central hypothesis:

For a design concept generation problem, an LLM provides more diverse design concepts when it is prompted with multiple professional personas using a sequential prompting strategy compared to (1) a parallel prompting strategy and (2) a collective prompting strategy.

Sub-hypotheses:

1a. Design concepts generated by LLMs operating with a virtual professional persona cover professional concepts that are semantically more aligned with concepts from the knowledge base of this persona.

1b. Topics in a professional knowledge base are semantically more aligned with topics within the same knowledge base, as compared to topics from other knowledge bases.

2. For some design concepts generated by an LLM prompted with some professional persona, a more diverse set of design concepts can be generated by prompting the LLM with a new professional persona to update these concepts.

3a. LLMs generate more diverse concepts when prompted with multiple personas using the parallel prompting strategy, compared to providing prompts without specifying any persona.

3b. Given multiple professional personas, LLMs generate more diverse design concepts when these personas are provided using the sequential prompting strategy, compared to the collective prompting strategy.

3c. Given multiple professional personas, LLMs generate more diverse design concepts when these personas are provided using the sequential prompting strategy, compared to the parallel prompting strategy.

The overview of the paper is as follows. First, a concise background of existing studies on design ideation, LLMs and design representations is presented. Second, the methods applied in this study are introduced, including (1) knowledge base construction, (2) design problem and professional persona selection and (3) prompting technique and strategy design. Third, each sub-hypothesis defined above is tested across seven design ideation problems using five different professional personas from various fields, and the results are discussed. Finally, the paper concludes by discussing the findings and their implications for engineering design.

2. Background

This section provides an overview of the study's relevant topics and reviews the existing literature. The discussion covers existing design ideation techniques, LLMs and their application to design ideation problems and design representations.

2.1. Design ideation

2.1.1. *Traditional design ideation techniques and limitations*

Design ideation consists of generating, developing and communicating ideas where an “idea” refers to a fundamental unit of thought that can take visual, tangible or abstract forms (Jonson 2005). According to Tschimmel (2012), there are usually two stages in a design activity: divergence and convergence. Designers can benefit from the diversity of ideas generated in the divergence stage to converge into more innovative design concepts. Different methods have been developed to support designers in generating solutions, such as brainstorming (Osborn 1957), morphological analysis (Álvarez & Ritchey 2015) and design heuristics (Daly *et al.* 2011).

Previous work has shown that the effectiveness of these methods depends on the creativity of human designers and can be restricted by a range of factors. For example, cognitive biases affect designers’ behavior in information interpretation, decision-making and reasoning (Hewstone, Rubin & Willis 2003; Boysen & Vogel 2009; Plews-Ogan *et al.* 2020). As a result, designers tend to favor certain design ideas while dismissing others (Agyemang, Andreae & McComb 2023). Groupthink is another factor that can restrict human creativity in the design ideation process, which was introduced by Janis (1982) as “a group where loyalty requires each member to avoid raising controversial issues.” This is the combined result of unconscious preferences for minimizing cognitive effort and a subconscious inclination to avoid social resistance, which can lead to low-quality decisions and hinder the group’s ability to explore a broader range of alternatives (Fox 2019; Akhmad, Chang & Deguchi 2021).

2.1.2. *Multidisciplinary collaboration in design ideation*

Given the cognitive and social constraints in traditional ideation methods, multidisciplinary collaboration has emerged as a valuable approach to overcoming such limitations and enriching the design process. According to Adams *et al.* (2009), multidisciplinary practice refers to the process of designers “joining together of disciplines to work on common problems and splitting apart when work is done.” Kleinsmann *et al.* (2012) also noted that collaboration requires crossing knowledge boundaries, emphasizing synthesis of practices and shared understanding.

In order to effectively boost multidisciplinary collaboration in design, several methods and strategies have been developed to support the communication and integration of diverse knowledge throughout this process. Van Helden *et al.* (2023) found that clearly defined rules for collaboration, division of labor and shared objectives significantly enhance collaborative effectiveness in student design teams. Several studies have also focused on how linguistic and cultural differences in communication among multidisciplinary design teams can influence team performance during collaborative design tasks (Jutraz & Zupancic 2014; D’Souza & Dastmalchi 2017).

In general, design teams made up of members from different disciplines tend to generate more diverse and creative design solutions, as designers bring different perspectives to the problem-solving process (Agogué *et al.* 2014; Gero, Yu & Wells 2019).

2.2. Large language models (LLMs)

LLMs, such as BERT (Devlin *et al.* 2019), Llama (Touvron *et al.* 2023), ChatGPT (Radford *et al.* 2018) and Gemini (Anil *et al.* 2023), have been increasingly popular in recent years. These models produce language that closely mimics writing created by humans since they are trained on extensive textual datasets, which enables them to answer questions and carry out a wide range of linguistic tasks (Kasneci *et al.* 2023). Higher performance in these tasks can be attained by applying fine-tuning approaches on smaller task-specific datasets (Wang *et al.* 2023). The state-of-the-art models outperform previous artificial intelligence (AI) models in general intelligence, effectively taking on novel and complex tasks such as machine vision, creative writing, coding and problem-solving in mathematics (Bubeck *et al.* 2023).

Among LLMs, the Generative Pre-training Transformer (GPT) has been iteratively improved from GPT-1 to GPT-5 (current at the time of this writing). Based on the transformer architecture, a GPT is trained using unsupervised learning methods on large text datasets. ChatGPT is a chatbot based on the GPT architecture. It is a potent tool with a wide range of applications across multiple areas (Liebrenz *et al.* 2023; McGee 2023; Thorp 2023; Wu *et al.* 2023).

2.2.1. LLM-assisted design

To tackle the limitations of human designer-based design techniques introduced in Section 2.1, various studies tested the feasibility of using LLMs in design tasks. For example, Zhu & Luo (2023) show that these models with certain customizations could generate verbal design concepts with a reasonable level of competence, while Wang *et al.* (2023) propose a prompt-guided method inspired by the function–behavior–structure model to progressively generate design ideas.

With the rapid development of LLMs, recent research has focused on integrating these models into various aspects of the design process. First, multiple studies have tested the capacity of LLMs to handle information and knowledge, from extracting and comprehending technical information from engineering documentation (Doris *et al.* 2024), to utilizing external knowledge to inform the design process (Han & Moghaddam 2024; Li, Ko & Ameri 2025). Second, multiple studies have focused on how LLMs can influence cognitive processes and education in design thinking (Agarwal, Jablokow & McComb 2024; Jiang, Huang & Shen 2024; Zhang, Zhao & Haddad 2024). Generally, these studies demonstrate that prompting LLMs with different cognitive styles can influence the feasibility and innovativeness of generated solutions, thereby enhancing design fluency, particularly for novice designers. Finally, LLMs have also been utilized to automate or support key aspects of the detailed design process, such as material selection (Grandi *et al.* 2025), requirements elicitation (Ataei *et al.* 2024), computer-aided design (CAD) object generation (Li, Sun & Sha 2025), engineering requirements tasks (Norheim *et al.* 2024) and additive manufacturing (Xie *et al.* 2025).

Researchers have also explored the performance of LLMs in the concept generation and ideation process. An early study by Ma *et al.* (2023) assessed the proficiency of LLMs in generating conceptual designs and contrasted the design ideas produced with crowd-sourced design concepts. The authors found that while LLMs can make more feasible and valuable design ideas, they are less novel and diverse. Their subsequent study tested the impact of parameter tuning and prompt variations on the diversity of design concepts generated by LLMs (Ma *et al.* 2025).

This study showed that there is no clear pattern between the parameters and concept diversity, and concept diversity is also highly responsive to the structure of the prompts provided. Chen *et al.* (2024) integrated LLMs with concept-knowledge (C-K) theory to support interdisciplinary knowledge retrieval and structured concept generation, demonstrating improved innovation and feasibility in design outcomes. Naghavi, Wang & Xu (2024) combined deep generative models with LLMs to reconstruct and generate porous metamaterial units, showing the potential of LLMs for advanced design representation and ideation. Given these results, introducing the concept of multidisciplinary collaboration into the LLM-assisted design process could potentially be beneficial in generating diverse and innovative design concepts.

2.2.2. Prompt engineering

For a language model, the prompt is the user-provided input that guides the model's output. The study of prompt engineering aims to create the best possible prompt inputs so that LLMs can produce the necessary (accurate or novel) responses. To effectively execute the specific tasks discussed in this study, it is essential to design prompts for LLMs that consider existing prompt engineering techniques.

Amatriain (2024) provided several pointers and strategies for the prompt design process, such as “encouraging the model to be factual,” “explicitly ending the prompt instructions,” and “being forceful.” Based on these instructions, more sophisticated methods for creating a prompt or set of prompts have been tried and documented. Wang, Moazeni & Klabjan (2025) proposed a Bayesian optimal learning framework for automated prompt engineering in LLMs, using a knowledge-gradient policy to efficiently discover high-performing prompts under limited evaluation budgets. Cheng *et al.* (2024) proposed a model-adaptive prompt optimization (MAPO) approach that adapts different LLMs using model-friendly prompts to enhance their capabilities across various downstream tasks. Yuksekogonul *et al.* (2024) introduced the TEXTGRAD framework, which enables automatic optimization of components in AI systems by backpropagating natural language feedback from LLMs based on a user-defined objective.

Specifically, various studies have examined how including personas in the prompts provided to LLMs can impact their performance in general tasks. These studies have led to different conclusions. Hu & Collier (2024) and Zheng *et al.* (2024) demonstrated that incorporating personas in prompts has minimal impact on the performance of LLMs in natural language processing (NLP)-related tasks. However, Luz de Araujo & Roth (2025) found that LLMs prompted with personas show greater variability in their responses to both objective and subjective tasks. It is worth noting that these studies have primarily focused on the impact of LLMs prompted by personas in NLP and question-answering tasks, while their impact on design-related problems remains underexplored.

2.3. Design representations

Representing design concepts is crucial for the design process. In addition to providing a strategy for efficient and systematic analysis and comparison of various design concepts, it also assists in determining the efficacy of different ideation

processes (Cash & Maier 2021). Over the years, various frameworks and methodologies have been developed to represent design concepts, which are utilized in different domains for diverse objectives. First, a design can be represented as a semantic network based on design descriptions in textual form (Sarica, Han & Luo 2023). This representation provides designers with insights into complex information in design concepts by utilizing an existing knowledge base to extract entities and relations from a textual design description. Furthermore, parametric design is an approach that defines a design concept using parameters that represent different attributes of the design concept, and the simple modification or optimization of each design is achieved through parameter changes (Monedero 2000). Parametric design is widely applied in areas such as CAD. Lastly, feature-based modeling attempts to illustrate how changes to each input feature impact the design as a whole by representing designs according to their features (Bhatt, Weller & Moura 2020). This method works especially well for engineering, production and assembly activities.

3. Methods

The approach adopted in this study aims to test how an LLM can be prompted to increase the diversity of design concepts it generates. In this approach, an LLM is first prompted with multiple professional personas to generate a set of concepts for a specific design problem. Different strategies for the prompting process are tested in this study, which are discussed later in this section. In each strategy, the same LLM is also prompted to summarize the design concepts generated into multiple professional terms. These terms are embedded within a space of professional knowledge bases to test their alignment with existing professional topics, and the diversity of a set of terms is also analyzed to assess the diversity of design concepts generated.

The specific steps include constructing professional knowledge bases, selecting design problems and professional personas, and designing prompting techniques and strategies. These steps are discussed in further detail in the following sections.

3.1. Knowledge base construction

In this paper, the knowledge base associated with a domain refers to the set of primary topics with which a typical expert within that domain would be familiar. For example, a trained mechanical engineer would be familiar with topics such as solid mechanics, fluid mechanics, machine design and heat transfer. In contrast, a trained computer scientist would be more familiar with topics such as AI, algorithm structures and human–computer interactions. These higher-level topics can be further decomposed into more specific topics such as the first law of thermodynamics, entropy and thermodynamic cycles for the mechanical engineering (ME) knowledge base, or NLP, search algorithms and user experience design for the computer science (CS) knowledge base.

In this study, knowledge bases were built for the following five disciplines: ME, chemical engineering (CHE), CS, biology (BIOL) and psychology (PSY), which correspond to the five professional personas selected for the prompting process. For each knowledge base, the following two-stage procedure was adopted to collect the topics: (1) core course collection and (2) topic extraction and preprocessing.

In the first step, the key courses for each field of study were determined by reviewing the curriculum of the associated undergraduate program. The courses selected were directly relevant to the specific topics within each major, so foundational courses such as calculus and physics were excluded. The selected courses ensured that a wide range of professional concepts was represented within each domain. In total, 45 courses were gathered across the five domains.

Subsequently, a standard textbook was acquired for each core course, and the table of contents from each textbook was used for topic extraction. Notably, the themes presented in the table of contents were chosen because they were more concise than other sources for extracting course topics, such as the course description or learning outcomes. Next, the topics for each course were extracted from the table of contents. The raw text data for the topics were cleaned up after this process. Headers such as “Introduction,” “Summary,” “Exercises” and “Overview,” which have general or ambiguous definitions, were excluded. All of the topics were transformed to lowercase because the text embedding API from OpenAI (Radford *et al.* 2023) used in this study is case-sensitive, except for proper nouns and acronyms, such as “SI units” and “DNA.” The creation of the knowledge base for this study was completed after the data-cleaning procedures. This knowledge base, divided into five separate domain-specific knowledge bases, comprises a total of 4469 topics. Table 1 provides a summary of the knowledge base.

3.2 Design problems and selection of professional personas

A set of criteria was used to select the design problems that are tested in this study from the design research literature. To encourage design ideas from designers with diverse professional backgrounds, the problem should first encompass a range of professional fields. Second, the problem statement should not contain any particular design restrictions that would prevent the generation of creative design concepts. Third, every design issue needs to be summarized into a single sentence to fit into the prompt provided to an LLM using a predetermined format.

The design ideation problems tested in this study were modified from the study by Goucher-Lambert & Cagan (2019) based on the criteria introduced above. These problems originated in earlier design literature and were all modified according to the same specifications for our investigation. An overview of the design ideation problems employed in this paper is provided in Table 2. As indicated in the previous section, five professional personas, namely mechanical

Table 1. Summary of the knowledge base across five professional domains		
Domain	Number of courses	Number of topics
Mechanical engineering	10	1705
Chemical engineering	10	1116
Computer science	6	795
Biology	9	399
Psychology	10	454
Total	45	4469

Table 2. Design problems used in this study

Problem No.	Problem description
1	A lightweight exercise device that can be used while traveling.
2	A device that can collect energy from human motion.
3	A device that disperses a light coating of a powdered substance over a surface.
4	A device that allows people to get a book that is out of reach.
5	An innovative product to froth milk.
6	A way to minimize accidents from people walking and texting on a cell phone.
7	A way to make drinking fountains accessible for all people.

engineer, chemical engineer, computer scientist, biologist and psychologist, were selected. These professional personas were chosen for two reasons: (a) they spanned a wide range of domains of study, from engineering to science and social science, and (b) they were in alignment with some undergraduate majors that are commonly offered in colleges.

3.3. Prompting techniques

3.3.1. Prompting with a single professional persona for design concept generation

This section introduces how LLMs were prompted with a professional persona to generate design concepts for a design problem. In this process, the design problem is provided to an LLM with a professional persona specified using the following prompt:

Prompt 1:

“I would like you to generate {num_solutions} different designs or solutions for {design_problem} using the knowledge base of a {profession}.

For each design concept, please provide a long paragraph of detailed technical description about how this design can be built and the specific professional concepts included in this design.”

When prompted to the LLM, “{num_solutions}” specifies the number of design concepts the LLM is expected to generate and “{design_problem}” is replaced by the problem string listed in Table 2 to specify the design problem. The “{profession}” part is replaced with one of the following strings to specify the professional persona LLM is asked to generate design concepts in: “mechanical engineer,” “chemical engineer,” “computer scientist,” “biologist” and “psychologist.”

With this prompt, a set of design concepts for the specified design problem is generated by the LLM with the specified professional persona.

3.3.2. Prompting with multiple professional personas for design concept generation

The following prompt was provided to an LLM to generate design concepts based on multiple professional personas through a single prompt:

Prompt 2:

“I would like you to generate {num_solutions} different designs or solutions for {design_problem} using the combined knowledge base of {professions}.

For each design concept, please provide a long paragraph of detailed technical description about how this design can be built and the specific professional concepts included in this design.”

This prompt is similar to the prompt defined for a single professional persona, with the only difference being that the “{professions}” part in the sentence above is replaced with the combined strings of the professional personas. The personas are separated using commas, such as “mechanical engineer, chemical engineer, computer scientist.” Using this prompt, a set of design concepts for the given design problem is generated by the LLM with the combination of specified personas.

3.3.3. Prompting with professional persona for design concept update

In the design ideation process, a new set of design concepts can be generated by updating an existing set of design concepts using the knowledge base of a new professional persona. This process is achieved by providing the following prompt to an LLM:

Prompt 3:

“Listed below are some design concepts for {design_problem}.

{design_solutions}

I would like you to generate {num_solutions} new designs or solutions for the same design problem by using the knowledge base of a {profession} to integrate or complete each of the design concepts generated above.

For each new design concept, please provide a long paragraph of detailed technical description about how this design can be built and the specific professional concepts included in this design.”

In this prompt, “{design_problem},” “{num_solutions}” and “{profession}” share the same meaning as in the first prompt. Also, the “{design_solutions}” part is replaced by the string of design concepts to be updated in this step. Using the prompt, a new set of design concepts is generated by applying the expertise of a new persona to update an existing set of design concepts.

3.3.4. Prompting for design concept summarization

The LLM-generated design concepts were summarized into a set of professional terms, and the alignments between these terms and topics from professional knowledge bases were evaluated to assess how professional knowledge in different fields is utilized in the LLM-generated design concepts. This process is

achieved by providing the following prompt to the same LLM used for design concept generation:

Prompt 4:

“Listed below are {num_solutions} design concepts for {design_problem}:

{design_solutions}

I would like you to summarize each design concept into {num_terms} professional terms strictly in the format of the following example:

Name: design concept.

Professional Terms: term 1; term 2; term 3; ...”

In this prompt, the “{num_solutions}” part specifies the number of design concepts to be summarized with this prompt; the “{design_problem}” part specifies the original design problem these design concepts correspond to; the “{design_solutions}” part is replaced by the string of design concepts to be summarized with this prompt; and the “{num_terms}” part specifies the number of professional terms each design concept is summarized into. With this prompt, the LLM is also instructed to formulate its output in the defined format, so that the terms it provides can be easily extracted later.

With this prompt, the design concepts are summarized into a set of professional terms that represent the professional knowledge involved.

3.4. Prompting strategies

This study adopts three different strategies to prompt LLMs with multiple professional personas for design concept generation: parallel prompting, collective prompting and sequential prompting. These correspond to the strategies defined in Hypothesis 3, and each strategy is implemented by applying the prompting techniques described above.

3.4.1. Parallel prompting

Parallel prompting refers to independently providing multiple prompts, each with a professional persona, to an LLM to generate design concepts for a design problem. As shown in Figure 1, the design problem is first presented independently to the LLM with different personas using Prompt 1 defined above, and multiple sets of design concepts are generated corresponding to the provided personas in this step. With this strategy, an LLM generates multiple sets of design concepts for a design ideation problem, each corresponding to a professional persona.

3.4.2. Collective prompting

In collective prompting, an LLM is provided with a single prompt that contains multiple professional personas for generating design concepts. This process begins with Prompt 2, as defined above, using the specified combination of personas. Only one set of design concepts is generated at this step. With this strategy, an LLM generates a single set of design concepts for a design ideation problem by combining the knowledge bases of different personas.

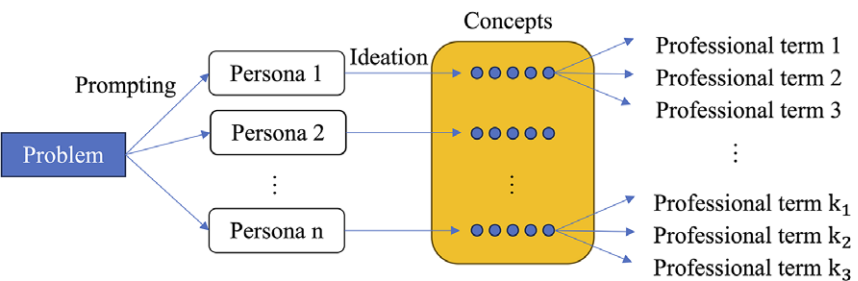


Figure 1. Parallel prompting strategy.

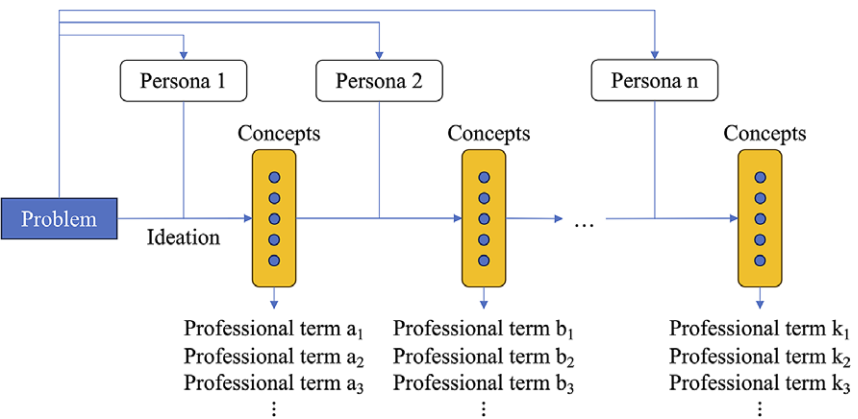


Figure 2. Sequential prompting strategy.

3.4.3. Sequential prompting

In sequential prompting, the LLM is provided with a sequence of prompts, each with a professional persona, to generate design concepts by iteratively updating existing design concepts. As shown in Figure 2, the LLM is first instructed to generate a set of design concepts using Prompt 1 with the first persona in the sequence. This set of design concepts is then provided to the same LLM with Prompt 2, where the LLM is prompted to generate a new set of design concepts for the same problem by updating the concepts provided using the knowledge base of the second persona in sequence. Subsequently, this update process is repeated using each of the remaining personas in the sequence, and a new set of design concepts is generated in each step by updating the concepts generated from the last step using the knowledge base of the persona specified in the current step. For a sequence of n professional personas, n sets of design concepts are generated in this process.

4. Experiments with design problems for the testing of hypotheses

This section presents experiments with selected design problems aimed at testing the hypotheses defined earlier in this article. Each hypothesis was tested using three different LLMs: ChatGPT (GPT-4o), Gemini-1.5 flash and Llama-3.1-8B-Instruct.

4.1. Evaluation of design concepts

Each design concept was represented by a set of related professional terms. The assessment of design concepts relied on the semantic analysis of associated terms, where semantic text analysis was achieved by numerically encoding each piece of textual data as a vector within an embedding space (Mikolov *et al.* 2013). Textual data were encoded into vectors using OpenAI’s “text-embedding-3-large” model, which maps text strings into a 3072-dimensional space. These semantic representations enabled the evaluation of similarity between two phrase groups and the analysis of the diversity of phrases within each group.

The *semantic similarity* of two phrases can be quantified by computing the cosine similarity of their vector representations in the embedding space (Rahutomo, Kitasuka & Aritsugi 2012). Table 3 provides an illustrative example of the cosine similarities for some phrase pairs in the embedding space of OpenAI’s “text-embedding-3-large” model. On this basis, the similarity between two groups of phrases can be measured by calculating the average pairwise cosine similarity between phrases in the first group and phrases in the second group.

The *diversity* of phrases within a group is calculated using the approach introduced by Ahmed & Fuge (2018). This method involved constructing a symmetric similarity matrix *L* for a collection of phrases *S*. Each element *L*_{*ij*} in the matrix quantifies the similarity between two phrases *i* and *j* within *S*. A value of *L*_{*ij*} = 1 indicates identical phrases, while *L*_{*ij*} = 0 denotes complete dissimilarity. The overall diversity of the phrases in *S* is determined with this matrix using the following equation:

$$Div(S) = (\det(L_S))^{1/n}.$$
 (1)

In this context, *n* denotes the number of phrases in set *S*. The approach assumes that the determinant of matrix *L* serves as an estimate of the extent of the high-dimensional space occupied by the vectors representing the phrases in *S*. A set with higher diversity among its phrases is expected to occupy a larger portion of this space, leading to a higher determinant value.

4.2. Experiments and evaluation metrics

This section provides a brief introduction to the experiments conducted to test each hypothesis in this study, as well as the evaluation of the results for each experiment.

Table 3. Cosine similarities for sample phrase pairs		
Phrase 1	Phrase 2	Cosine similarity
cloud computing service	financial risk assessment	0.138
cloud computing service	renewable energy solutions	0.212
cloud computing service	machine learning model	0.331
cloud computing service	cloud infrastructure deployment	0.667

Hypothesis 1a is on the alignment between design concepts generated by the LLMs with a specific persona and different professional knowledge bases. In contrast, Hypothesis 1b examines the difference between the semantic alignment of topics within the same knowledge base and that of topics across different knowledge bases. To test Hypothesis 1a, five sets of design concepts were generated for each design problem using the parallel prompting strategy with the five selected personas. A set of related professional terms then summarizes each set of concepts. The similarity between a set of terms and each professional knowledge base is measured using the group similarity metric introduced above. The testing of Hypothesis 1b includes two steps. For topics from the same knowledge base, semantic similarity is captured by the average intragroup pairwise cosine similarity within these topics. In contrast, semantic similarity between two different knowledge bases is represented using the group similarity metric.

Hypothesis 2 assesses whether prompting LLMs to update an existing set of design concepts with a professional persona can generate a more diverse set of design concepts. In this experiment, an LLM is first prompted with one persona to generate a set of design concepts, which are then revised under a different persona to produce a new set of concepts. Each set of concepts is then summarized into a set of professional terms, and the diversity of each set of terms is analyzed. Finally, the diversity scores of both sets of terms are compared to capture the change in design concept diversity through the update process.

Hypotheses 3a–3c are formulated to test the difference in the diversity of design concepts generated by LLMs when prompted using different strategies, and two strategies are compared explicitly in each hypothesis. In the experiment for each hypothesis, two sets of design concepts are generated using the specified strategies, and each set of concepts is summarized using related professional terms. The diversity scores of both sets of terms are then calculated and compared to demonstrate how the diversity of design concepts generated by the LLM differs when prompted with different strategies.

4.3. Sample outputs by LLM

4.3.1. Sample design concept generation and summarization

To illustrate the process of LLM-assisted design concept generation, we provide a sample design problem to ChatGPT (GPT-4o) using Prompt 1 with some professional personas to generate multiple sets of design concepts. Each set is then provided back to ChatGPT (GPT-4o) with Prompt 4 and summarized in professional terms. In this case, the design problem provided is “an innovative product to froth milk,” and Prompt 1 is modified accordingly as:

“I would like you to generate 5 different designs or solutions for an innovative product to froth milk using the knowledge base of a {profession}.”

For each design concept, please provide a long paragraph of detailed technical description about how this design can be built and the specific professional concepts included in this design.”

The “{profession}” part in the prompt was replaced to specify the persona used in this concept generation process. Each design concept generated at this step was

then summarized into 10 professional terms by modifying Prompt 4 in the following format:

“Listed below are 5 design concepts for an innovative product to froth milk:

{design_concepts}

I would like you to summarize each design concept into 10 professional terms strictly in the format of the following example:

Name: design concept.

Professional Terms: term 1; term 2; term 3; ...”

The “{design_concepts}” part in the prompt was replaced by the design concepts generated by ChatGPT (GPT-4o) in the previous step. With these two prompts, five design concepts are generated for the selected design problem using a specified persona, and each design concept is summarized into five professional terms. For illustration, the output design concepts and summarized terms using the mechanical engineer and psychologist personas are provided in [Tables 4](#) and [5](#), respectively.

Table 4. Design concepts and summarized terms generated for the selected problem using ChatGPT (GPT-4o) with the mechanical engineer persona		
Concept no.	Name	Summarized terms
1	Electromagnetic Frother	<ul style="list-style-type: none">• electromagnetic induction• eddy currents• magnetically coupled impeller• microcontroller control• thermal dynamics
2	Ultrasonic Frother	<ul style="list-style-type: none">• ultrasonic transducer• cavitation bubbles• high-frequency vibrations• sound insulation• temperature sensor
3	Steam-Injection Frother	<ul style="list-style-type: none">• high-efficiency boiler• thermally insulated tube• steam wand• pressure regulator• thermostatic valve
4	Centrifugal Frother	<ul style="list-style-type: none">• centrifugal force• high-speed motor• frothing disk• vortex motion• protective casing
5	Aerator-Pump Frother	<ul style="list-style-type: none">• diaphragm pump• aeration nozzle• micro-holes• miniature electric motor• syncing microcontroller

Table 5. Design concepts and summarized terms generated for the selected problem using ChatGPT (GPT-4o) with the psychologist persona

Concept no.	Name	Summarized terms
1	Ergonomic Handheld Milk Frother	<ul style="list-style-type: none">• ergonomics• vibration reduction• sensory feedback• non-slip grip• comfort
2	Mindful Frothing Station	<ul style="list-style-type: none">• mindfulness• ambient lighting• minimal noise• anxiety reduction• calming sound
3	Social Connectivity Frother	<ul style="list-style-type: none">• Wi-Fi connectivity• recipe sharing• live broadcasting• virtual interaction• community building
4	Personalized Learning Frother	<ul style="list-style-type: none">• AI coaching• real-time feedback• performance tracking• gamification• skill enhancement
5	Memory-Infused Nostalgia Frother	<ul style="list-style-type: none">• customizable scents• sensory experience• aromatherapy• nostalgic music• positive mood

4.3.2. Sample design concepts generated by LLM using sequential prompting strategy

This section provides some sample design concepts provided by ChatGPT (GPT-4o) prompted using the sequential prompting strategy. Specifically, the design problem selected for the generation process is also “an innovative product to froth milk,” and the tested sequence is: “**computer scientist** → **biologist** → **psychologist** → **mechanical engineer** → **chemical engineer**.” The sequential prompting process includes five steps, where the LLM is provided with one persona in each step. The LLM generates five design concepts in each step, and these concepts are summarized into 25 professional terms, resulting in a total of 125 summarized terms across five design concepts. The generated design concepts for the selected problem using the selected sequence are provided in Table 6.

According to Table 6, when prompted with the first persona in the sequence, which is “computer scientist,” ChatGPT (GPT-4o) generates design concepts that are related to the professional knowledge of a computer scientist, such as AI, IoT, robotics, augmented reality and voice control. For each of the subsequent steps, ChatGPT (GPT-4o) updates each existing design concept using the professional knowledge of the persona prompted in the current step. For example, updating the

Table 6. Sample design concepts generated by ChatGPT (GPT-4o) using a selected sequence of personas for the design problem “an innovative product to froth milk”

Step	Persona	Design concepts generated by ChatGPT (GPT-4o)
1	Computer scientist	<ul style="list-style-type: none">• AI-Assisted Frother• IoT-Enabled Smart Frother• Robotic Arm Frother• Augmented Reality Froth Assister• Voice-Controlled Cloud Frother
2	Biologist	<ul style="list-style-type: none">• Bio-Inspired AI Frother• Eco-Friendly IoT Frother• Biomechanical Frother• Symbiotic AR Froth Assister• Sensory-Driven Voice-Controlled Frother
3	Psychologist	<ul style="list-style-type: none">• Cognitive AI Frother• Behaviorally-Inspired Eco-Friendly IoT Frother• Experience-Oriented Biomechanical Frother• Communal AR Froth Assister• Emotionally Intelligent Sensory-Driven Voice-Controlled Frother
4	Mechanical engineer	<ul style="list-style-type: none">• Adaptive Gear-Controlled Cognitive AI Frother• Renewable Energy-Powered Behaviorally-Inspired Eco-Friendly IoT Frother• Ergonomically Enhanced Experience-Oriented Biomechanical Frother• Collaborative Modular Communal AR Froth Assister• Emotionally Intelligent Sensory-Driven Hydraulic Voice-Controlled Frother
5	Chemical engineer	<ul style="list-style-type: none">• Nano-Emulsion AI Enhanced Frother• Bio-Catalytic Eco-Friendly IoT Frother• Viscoelastic Material Ergonomic Frother• Phase-Change Material Enhanced AR Froth Module• Electrolyte-Infused Hydraulic System

design concept “AI-Assisted Frother” using the persona “biologist” leads to “Bio-Inspired Frother,” and updating this concept using “psychologist,” “mechanical engineer” and “chemical engineer” personas, respectively, provide the following design concepts: “Cognitive AI Frother,” “Adaptive Gear-Controlled Cognitive AI Frother” and “Nano-Emulsion AI Enhanced Frother.” Generally speaking, when prompted with the sequential prompting strategy, ChatGPT (GPT-4o) is capable of generating new design concepts and updating existing design concepts by leveraging the professional knowledge of each persona in the sequence.

4.4. Hypothesis testing results

4.4.1. Results for Hypothesis 1a

Hypothesis 1a compares the semantic alignment between LLM-generated concepts and professional knowledge bases. The results for all three LLMs are shown in Table 7, where the results for each LLM are computed across all seven design problems. In each matrix, each row corresponds to the persona provided to the

Table 7. Group similarity between LLM-generated concepts and professional knowledge bases (results for Hypothesis 1a)

ChatGPT					
	ME	CHE	CS	BIOL	PSY
ME	0.197	0.179	0.158	0.134	0.133
CHE	0.193	0.183	0.152	0.142	0.130
CS	0.169	0.163	0.172	0.135	0.144
BIOL	0.180	0.170	0.152	0.152	0.142
PSY	0.162	0.157	0.157	0.134	0.164

Gemini					
	ME	CHE	CS	BIOL	PSY
ME	0.193	0.178	0.158	0.144	0.156
CHE	0.185	0.173	0.153	0.143	0.148
CS	0.182	0.173	0.164	0.145	0.158
BIOL	0.177	0.169	0.153	0.157	0.154
PSY	0.155	0.152	0.156	0.140	0.179

Llama					
	ME	CHE	CS	BIOL	PSY
ME	0.200	0.189	0.161	0.137	0.141
CHE	0.197	0.191	0.159	0.138	0.142
CS	0.183	0.180	0.162	0.131	0.144
BIOL	0.182	0.178	0.161	0.145	0.154
PSY	0.158	0.158	0.156	0.134	0.174

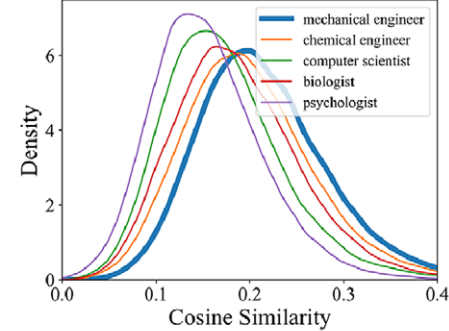
0.2

0.13

LLM in the concept generation process, while each column corresponds to the professional knowledge base with which the generated concepts are compared. It is observed that for a specific professional knowledge base, the concepts generated by an LLM with the profession corresponding to that knowledge base yield a higher average cosine similarity with the topics in this knowledge base, compared to concepts generated by an LLM with other professional personas. This suggests that when operating with a virtual professional persona, an LLM can generate design concepts that leverage more professional knowledge associated with this persona.

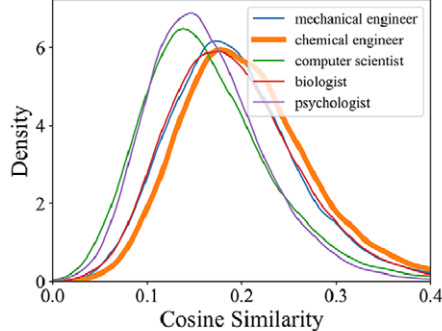
Figures 3 and 4 show the sample results of this alignment between ChatGPT (GPT-4o)-generated design concepts for the design problem “an innovative product to froth milk” and the five professional knowledge bases. Specifically, each subfigure in Figure 3 shows the probability density of pairwise semantic alignments between professional terms summarized from ChatGPT-generated concepts with different professional personas and topics from a knowledge base, while each subfigure in Figure 4 shows the cumulative distribution of the alignments. In Figure 3, a distribution curve with a higher mean value represents a better alignment of the ChatGPT (GPT-4o) response and the knowledge base. In contrast, a curve with a lower mean represents a lower alignment. However, in Figure 4, a better alignment is represented by the curve that reaches its peak later than others.

Problem 5, Mechanical Engineering Knowledge Base



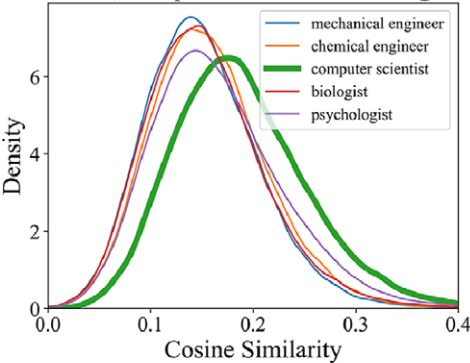
(a) Mechanical engineering knowledge base

Problem 5, Chemical Engineering Knowledge Base



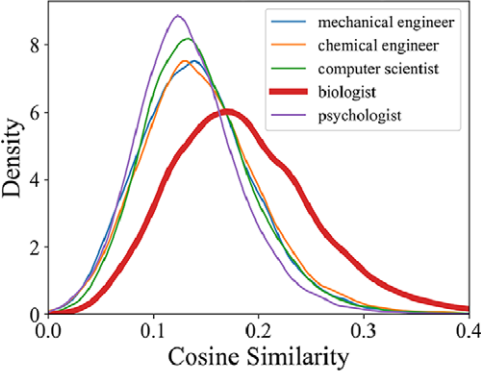
(b) Chemical engineering knowledge base

Problem 5, Computer Scientist Knowledge Base



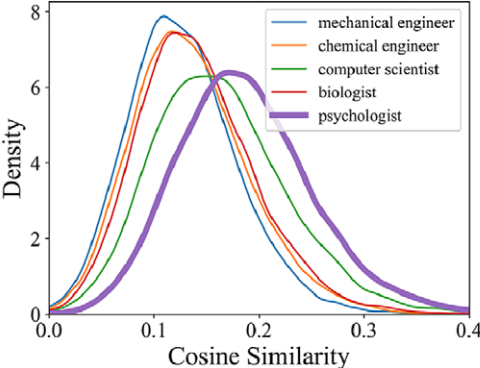
(c) Computer science knowledge base

Problem 5, Biologist Knowledge Base



(d) Biology knowledge base

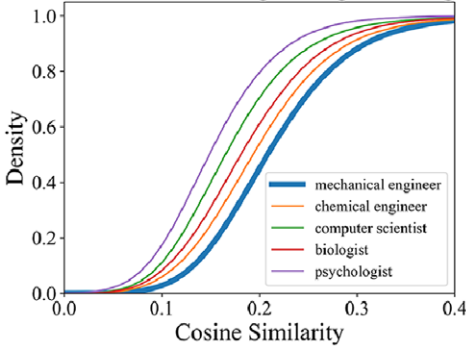
Problem 5, Psychologist Knowledge Base



(e) Psychology knowledge base

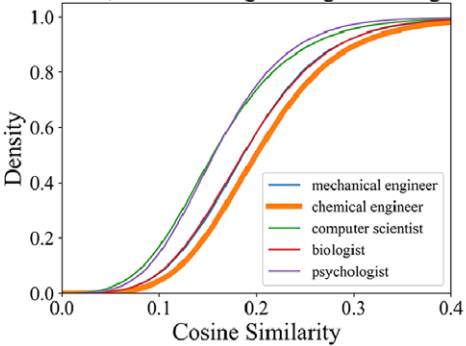
Figure 3. Probability density functions of ChatGPT (GPT-4o)-generated sample design concepts for “an innovative product to froth milk,” compared with different professional knowledge bases.

Problem 5, Mechanical Engineering Knowledge Base



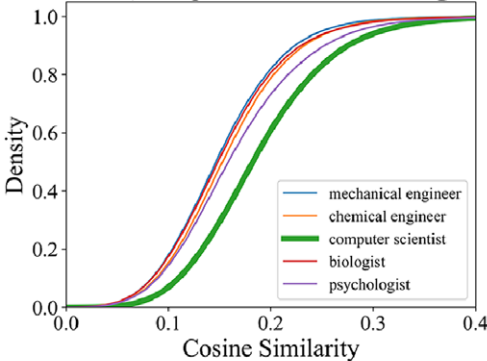
(a) Mechanical engineering knowledge base

Problem 5, Chemical Engineering Knowledge Base



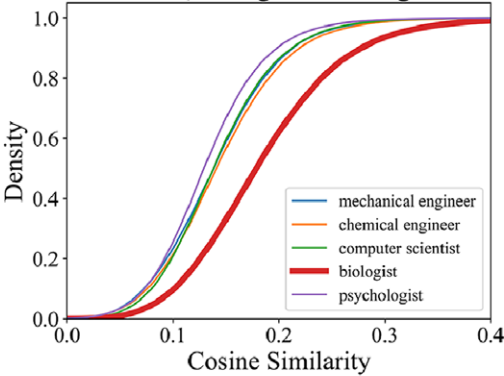
(b) Chemical engineering knowledge base

Problem 5, Computer Scientist Knowledge Base



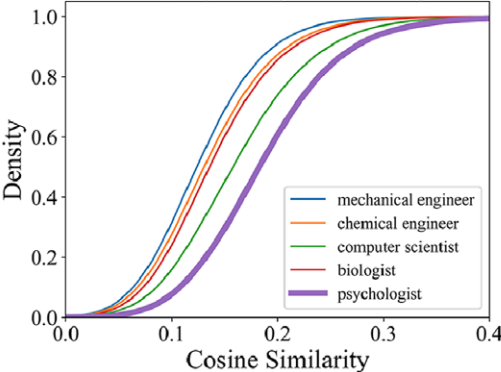
(c) Computer science knowledge base

Problem 5, Biologist Knowledge Base



(d) Biology knowledge base

Problem 5, Psychologist Knowledge Base



(e) Psychology knowledge base

Figure 4. Cumulative density functions of ChatGPT (GPT-4o)-generated sample design concepts for “an innovative product to froth milk,” compared with different professional knowledge bases.

Table 8. Group similarity between different professional knowledge bases (results for Hypothesis 1b)

	ME	CHE	CS	BIOL	PSY
ME	0.223	0.212	0.165	0.129	0.137
CHE	0.212	0.233	0.172	0.138	0.149
CS	0.165	0.172	0.220	0.135	0.151
BIOL	0.129	0.138	0.135	0.234	0.138
PSY	0.137	0.149	0.151	0.138	0.207

According to [Figures 3](#) and [4](#), for topics within a specific professional knowledge base, the design concepts generated by ChatGPT (GPT-4o) with the corresponding professional persona have a dominantly higher average similarity with these topics, as compared to concepts generated by ChatGPT (GPT-4o) with other personas. This shows that ChatGPT (GPT-4o), when prompted with a specific professional persona, is more capable of generating design concepts by leveraging the topics from the corresponding knowledge base than when prompted with other personas.

4.4.2. Results for Hypothesis 1b

Hypothesis 1b was tested by calculating the average intragroup pairwise cosine similarity within each knowledge base and the group similarity between two different knowledge bases. According to [Table 8](#), values in diagonal cells are generally higher than values in off-diagonal cells. This reflects that topics within a professional knowledge base are semantically more aligned with topics from the same knowledge base, as compared to topics from a different knowledge base. Thus, we can infer from this result that topics from different knowledge bases are semantically diverse, which confirms Hypothesis 1b.

4.4.3. Results for Hypothesis 2

Hypothesis 2 evaluates how updating existing design concepts using an LLM with a professional persona would impact the diversity of design concepts. In this experiment, the LLM first generates five sets of design concepts for a design problem using the parallel prompting strategy with the five personas in this study, namely S_i , where i refers to the persona used for concept generation. Next, each S_i is updated using Prompt 3 with the other four personas, respectively, to generate four new sets of design concepts, denoted as $S_{i,j}(i \neq j)$, where j refers to the persona used in the update process. Each of S_i and $S_{i,j}$ is then summarized by LLM using Prompt 4 into 50 related professional terms, and a diversity score is calculated to reflect the semantic diversity of terms involved in each set, denoted as D_i and $D_{i,j}$. Finally, each score $D_{i,j}$ is compared to the corresponding D_i to reflect the change in the diversity of LLM-generated design concepts throughout the update process. This change is represented as a percentage change in each cell of [Table 9](#), where a positive entry denotes an increase in diversity, while a negative entry denotes a decrease.

[Table 9](#) presents the results for all three LLMs, where each cell shows the percentage change between each $D_{i,j}$ and its corresponding D_i . In this table,

Table 9. Experiment results for Hypothesis 2

		Iteration 2				
		ME	CHE	CS	BIOL	PSY
Iteration 1	ME		0.97%	-1.24%	0.09%	1.52%
	CHE	-0.72%		0.84%	-0.68%	2.19%
	CS	-2.59%	-2.99%		-1.14%	-0.67%
	BIOL	-2.00%	-1.22%	-3.24%		0.38%
	PSY	-4.08%	-3.60%	-3.55%	-0.95%	

(a) ChatGPT (GPT-4o)

		Iteration 2				
		ME	CHE	CS	BIOL	PSY
Iteration 1	ME		5.35%	5.43%	4.33%	3.23%
	CHE	6.74%		11.09%	8.42%	13.90%
	CS	-0.25%	9.13%		4.71%	7.71%
	BIOL	-3.33%	4.21%	4.73%		2.17%
	PSY	2.67%	7.23%	9.44%	7.65%	

(b) Gemini-1.5 flash

		Iteration 2				
		ME	CHE	CS	BIOL	PSY
Iteration 1	ME		-6.27%	-6.20%	-7.08%	2.45%
	CHE	0.83%		-4.13%	-2.73%	4.22%
	CS	-10.28%	-8.44%		-8.43%	-7.46%
	BIOL	-3.78%	-9.29%	5.25%		-0.25%
	PSY	-4.75%	-5.00%	0.05%	0.02%	

(c) Llama-3.1-8B-Instruct

the cells with a positive value are highlighted in pink, indicating an increase in the diversity of design concepts. According to Table 9, updating existing design concepts using Gemini-1.5 flash with a persona would yield design concepts with increased diversity, while updating using ChatGPT (GPT-4o) or Llama-3.1-8B-Instruct would lead to no significant increase in the diversity of design concepts.

4.4.4. Results for Hypothesis 3a

Hypothesis 3a compares the diversity of design concepts generated by an LLM when it is (1) prompted with no professional persona and (2) prompted with multiple professional personas using parallel prompting. Specifically, the LLM is first prompted with no persona to generate 25 design concepts, and each concept is then summarized into five professional terms, providing a total of 125 terms. On the contrary, the same LLM is also prompted to generate five design concepts for each of the five personas in this study using the parallel prompting strategy. Each concept is summarized into five professional terms, providing a total of 25 design concepts and 125 terms. A comparison of diversity for both sets of terms generated by each LLM is listed in Figure 5. It is observed that for each LLM, the design concepts generated when prompted in parallel with multiple personas yield a better diversity score than when prompted with no persona.

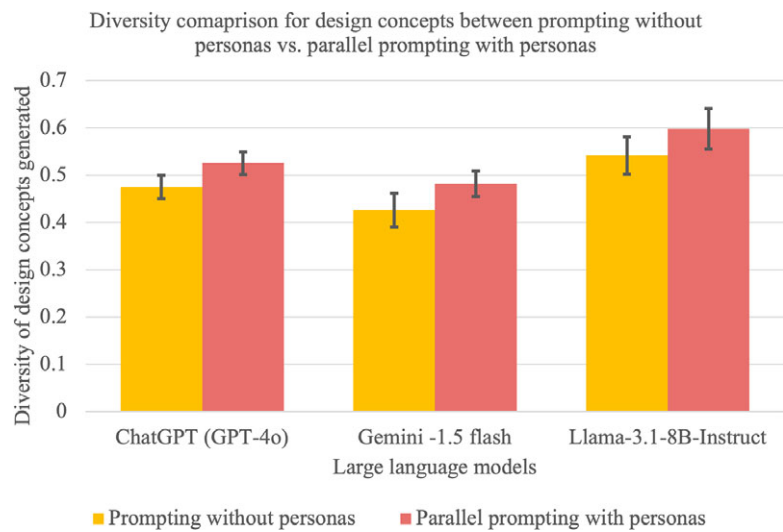


Figure 5. Experiment results for Hypothesis 3a.

4.4.5. Results for Hypotheses 3b and 3c

Hypotheses 3b and 3c compare the diversity of design concepts generated by an LLM when prompted with collective prompting, parallel prompting and sequential prompting strategies, respectively. In this experiment, the LLM is prompted with the five personas to generate 25 design concepts using each strategy defined above, and each set of concepts is summarized into 125 professional terms. The diversity score of each set of terms is then computed and compared. Notably, since different sequences of personas can lead to various design concepts, five sequences are tested for the sequential prompting strategy. Both the results of all sequences and the sequence that leads to design concepts with the highest diversity score are recorded in each experiment. The sequences tested in this study are:

- Sequence 1: ME → CHE → CS → BIOL → PSY
- Sequence 2: CHE → CS → BIOL → PSY → ME
- Sequence 3: CS → BIOL → PSY → ME → CHE
- Sequence 4: BIOL → PSY → ME → CHE → CS
- Sequence 5: PSY → ME → CHE → CS → BIOL

The results of all strategies are shown in Figure 6, where each subfigure represents the comparison using a different LLM. According to Figure 6, collective prompting generates design concepts with the lowest diversity among the three strategies. Moreover, for each LLM, parallel prompting overall provides design concepts with a level of diversity similar to that of concepts generated using the sequential prompting strategy with the “optimal” sequence of personas, where the “optimal” sequence is obtained by picking the sequence that leads to the most diverse design concepts out of all sequences tested.

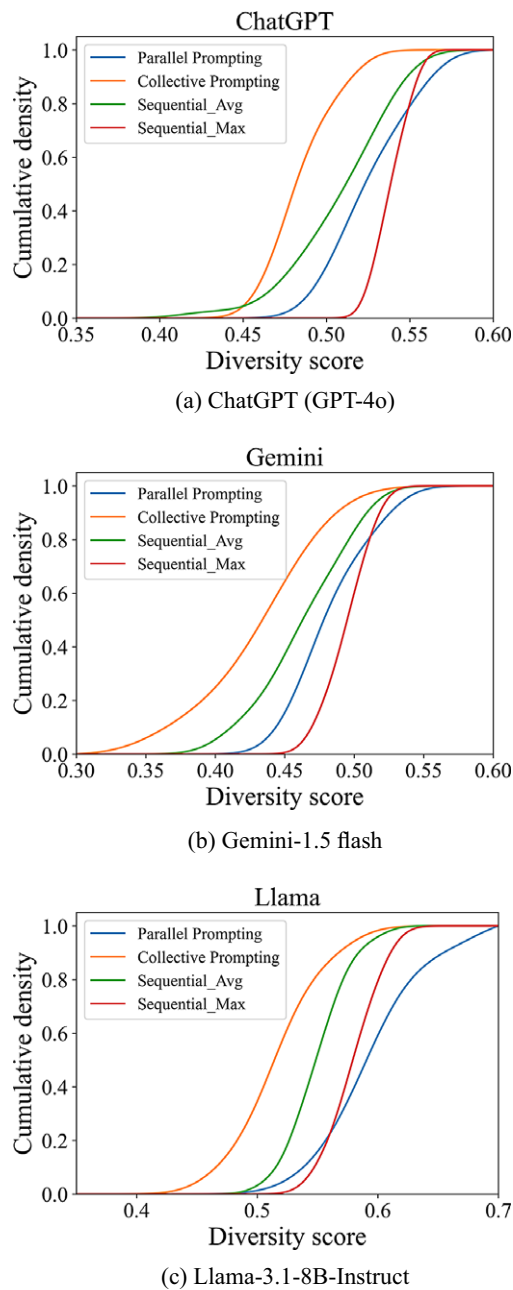


Figure 6. Experiment results for Hypotheses 3b and 3c.

4.4.6. Further analysis of sequential prompting strategy

For the sequential prompting strategy, one key aspect to explore is the underlying mechanism by which it affects the diversity level of the design concepts generated. To investigate this, we conduct further analysis of the concepts generated during the sequential prompting process.

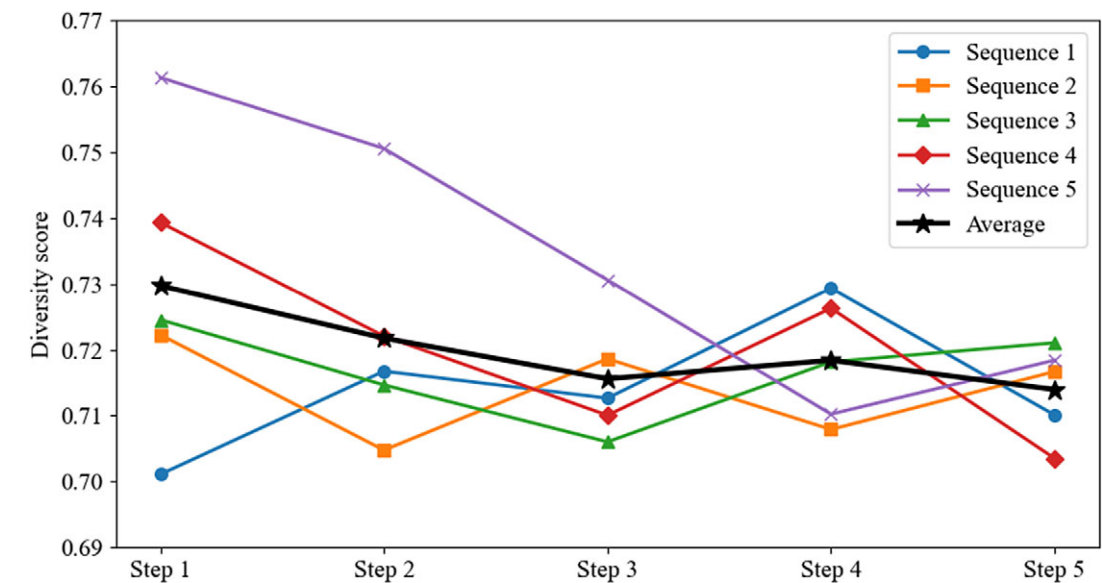


Figure 7. Diversity scores of design concepts generated by ChatGPT (GPT-4o) at each step of the sequential prompting process for the selected design problem.

First, we aim to capture how the diversity of design concepts generated at each step changes with the introduction of a new persona. For the five sequences of personas and the same design problem defined in Section 4.4.5, we measure the diversity score of the design concepts generated at each step of the sequential prompting process. The results are shown in Figure 7. It is observed that for the selected design problem, the diversity score of the design concepts generated at each step of the sequential prompting process remains relatively stable, indicating that introducing a new persona at each step does not significantly impact the diversity level of the generated design concepts.

We also investigate how the LLM leverages different professional knowledge bases throughout the sequential process, as reflected in the alignment between the design concepts generated at each step of the sequential prompting process, and each professional knowledge base. We use the five sequences of personas and the same design problem as in the previous experiment to measure the similarity (using the method introduced in Section 4.1) between the design concepts generated at each step and each professional knowledge base. The results are shown in Figure 8. This figure presents a comparison of all five sequences tested in this study, where each entry in the table represents the similarity between the design concepts generated at a specific step of the sequence corresponding to the current table and a professional knowledge base. Specifically, the boxed cells indicate the comparison between design concepts generated using a persona and the professional knowledge base corresponding to that persona. For example, Sequence 3 is CS → BIOL → PSY → ME → CHE. At Step 1 in Sequence 3, the boxed cell represents the comparison between the design concepts generated and the professional knowledge base of CS, the first persona in this sequence.

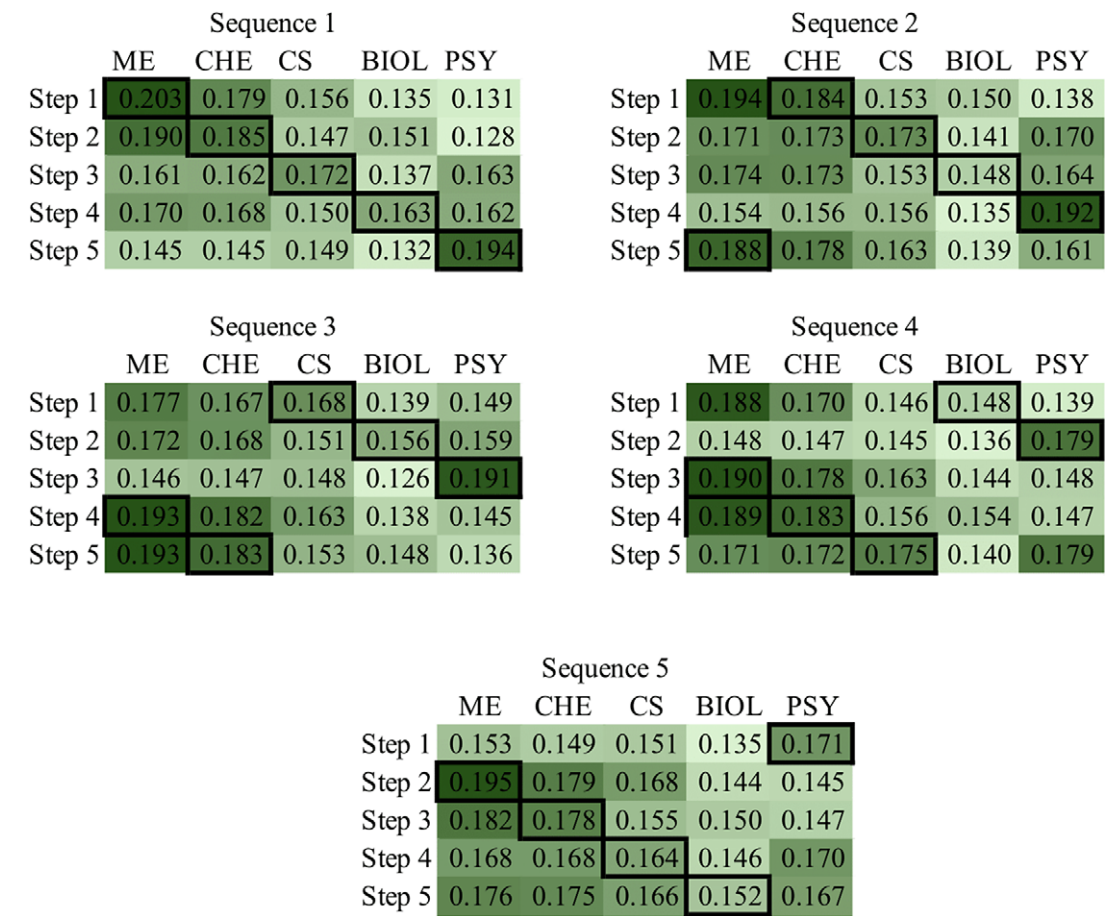


Figure 8. Group similarity between LLM-generated design concepts at each step of the sequential prompting process and professional knowledge bases.

As shown in Figure 8, at each step of the sequential prompting process, the LLM generates design concepts that are more similar to the professional knowledge base of the persona provided at the current step, as compared to the professional knowledge base of other personas. For example, for Sequence 3, the design concepts generated at the first step are more aligned with the professional knowledge base of CS, compared to any of the subsequent steps. At Step 2, they align more with BIOL, at Step 3 with PSY, and so on. Furthermore, the results also show that for any of the later steps in the sequential prompting process, the design concepts generated are not significantly aligned with the professional knowledge base of previously used personas in the sequence. For example, the third persona used in Sequence 3 is psychologist (PSY), with computer scientist (CS) and biologist (BIOL) being the first two personas. However, according to the results for Step 3 in the third table, the design concepts generated are more aligned with the professional knowledge base of PSY, but not significantly more aligned with the professional knowledge bases of CS and BIOL. These results together indicate that, when prompted with multiple personas using the sequential prompting strategy,

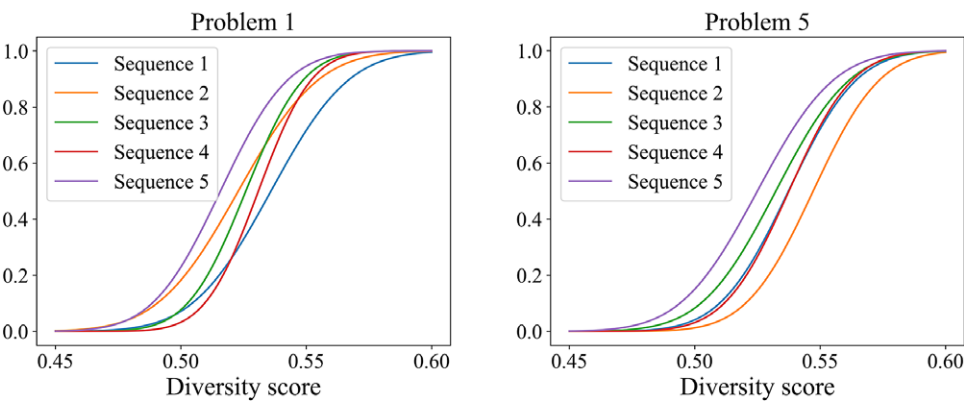


Figure 9. Cumulative density functions of diversity scores of design concepts generated by ChatGPT (GPT-4o) using different sequences of personas for selected design problems.

Table 10. Mean and standard deviation of diversity scores of design concepts generated by ChatGPT (GPT-4o) using different sequences of personas for selected design problems

	Problem 1		Problem 5	
	Mean	Standard deviation	Mean	Standard deviation
Sequence 1	0.536	0.025	0.537	0.021
Sequence 2	0.523	0.025	0.547	0.021
Sequence 3	0.526	0.018	0.532	0.023
Sequence 4	0.531	0.017	0.538	0.020
Sequence 5	0.516	0.021	0.525	0.024

LLM generates design concepts at each step using the professional knowledge base of the persona of the current step alone, instead of integrating the professional knowledge base of this persona with the design concepts generated at preceding steps of the sequence.

The experiment results from [Section 4.4.5](#) show that LLMs prompted with sequential and parallel prompting strategies generate design concepts with similar diversity scores. On this basis, we also explore how the diversity of design concepts differs when the LLM is prompted using the sequential prompting strategy with different sequences of professional personas. We again use the sequential prompting strategy with the same five sequences of personas as in the previous experiment to prompt ChatGPT (GPT-4o) to generate design concepts for design Problems 1 and 5. The results are shown in [Figure 9](#) and [Table 10](#). [Figure 9](#) shows the distribution of diversity scores of design concepts generated by ChatGPT (GPT-4o) using each sequence of personas for the selected design problems, while the numerical results for mean and standard deviation of each distribution are shown in [Table 10](#).

It is observed that for the selected design problems, the design concepts generated by ChatGPT (GPT-4o) share a similar level of diversity when prompted

with the five tested sequences of personas. Specifically, Sequence 1 yields the most diverse design concepts for Problem 1, Sequence 2 yields the most diverse design concepts for Problem 5, and Sequence 5 yields the least diverse design concepts for both design problems.

5. Conclusion and future work

This paper provides insights into how the diversity of design concepts generated by language models can be increased by prompting LLMs with virtual professional personas. Specifically, multiple hypotheses are formulated to compare different strategies for prompting LLMs with professional personas to generate design concepts. The methods used to test the hypotheses include knowledge base construction, selection of design problems and professional personas, and prompt design. Among all the prompting strategies studied in this paper, the results reveal that LLMs generate more diverse design concepts when prompted with the parallel prompting strategy and the sequential prompting strategy (with the best sequence), as compared to when prompted with the collective prompting strategy. We also conduct further analysis on the sequential prompting strategy, and the results reveal that (1) design concepts generated by an LLM share a similar level of diversity when it is prompted with different sequences of personas; and (2) LLM, when prompted with multiple personas using the sequential prompting strategy, generates design concepts at each step mostly using the professional knowledge base of the persona of the current step alone.

Using the strategies studied in this paper, designers can address the design concept generation problems with professional topics from the combined knowledge base of multiple professional personas. By prompting LLMs using the two strategies defined above, the diversity of design concepts generated can be increased during the divergence stage of design, which is beneficial for designers to yield more innovative design concepts.

This article lays the groundwork for further exploration of the application of LLMs with professional personas in the design process. First, while this study evaluates five personas using multiple sample sequences within a sequential prompting strategy, future research could extend this approach to a broader range of professional personas. Additionally, we used phrases such as “mechanical engineer” to define a persona. In the future, richer descriptions could be developed by incorporating more detailed attributes – such as specific knowledge and expertise in a professional domain, and their associated cognitive styles – to create more nuanced personas. Finally, human studies could be conducted to investigate how these persona-based strategies influence the creativity and diversity of ideas generated by human designers during the ideation process.

References

- Adams, R. S., Chen, J. C., Davis, K. M. & McKenna, K. J. 2009 Exploring the boundaries: integrating mathematical thinking into design reasoning in interdisciplinary design projects. In *About: Designing. analysing Design Meetings* (ed. J. McDonnell & P. Lloyd), pp. 173–198. CRC Press.

- Agarwal, V., Jablokow, K. & McComb, C. 2024 Putting the ghost in the machine: emulating cognitive style in large language models. *Journal of Computing and Information Science in Engineering* 25 (2), 021002; doi:[10.1115/1.4066857](https://doi.org/10.1115/1.4066857).
- Agogu , M., Poir , N., Pineau, A., Houd , O. & Cassotti, M. 2014 The impact of age and training on creativity: a design-theory approach to study fixation effects. *Thinking Skills and Creativity* 11, 33–41; doi:[10.1016/j.tsc.2013.10.002](https://doi.org/10.1016/j.tsc.2013.10.002).
- Agyemang, M., Andreae, D. A. & McComb, C. 2023 Uncovering potential bias in engineering design: a comparative review of bias research in medicine. *Design Science* 9, e17; doi:[10.1017/dsj.2023.17](https://doi.org/10.1017/dsj.2023.17).
- Ahmed, F. & Fuge, M. 2018 Ranking ideas for diversity and quality. *Journal of Mechanical Design* 140 (11), 111405; doi:[10.1115/1.4040641](https://doi.org/10.1115/1.4040641).
- Akhmad, M., Chang, S. & Deguchi, H. 2021 Closed-mindedness and insulation in groupthink: their effects and the devil’s advocacy as a preventive measure. *Journal of Computational Social Science* 4, 455–478; doi:[10.1007/s42001-020-00083-8](https://doi.org/10.1007/s42001-020-00083-8).
-  lvarez, A. & Ritchey, T. 2015 Applications of general morphological analysis: from engineering design to policy analysis. *Acta Morphologica Generalis* 4 (1), 1–22. <https://www.amg.swemorph.com>.
- Amatriain, X. 2024 Prompt design and engineering: introduction and advanced methods. Preprint. <https://arxiv.org/abs/2401.14423v2>.
- Anil, R., Borgeaud, S., Wu, Y., et al. 2023 Gemini: a family of highly capable multimodal models. Preprint. <http://arxiv.org/abs/2312.11805>.
- Ataei, M., Cheong, H., Grandi, D., Wang, Y., Morris, N. & Tessier, A. 2024 Elictron: an LLM agent-based simulation framework for design requirements elicitation. Preprint. <https://arxiv.org/abs/2404.16045>.
- Bhatt, U., Weller, A. & Moura, J. M. F. 2020 Evaluating and aggregating feature-based model explanations. Preprint. <https://arxiv.org/abs/2005.00631>.
- Boysen, G. A. & Vogel, D. L. 2009 Bias in the classroom: types, frequencies, and responses. *Teaching of Psychology* 36 (1), 12–17; doi:[10.1080/00986280802529038](https://doi.org/10.1080/00986280802529038).
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T. & Zhang, Y. 2023 Sparks of artificial general intelligence: early experiments with GPT-4. Preprint. <https://arxiv.org/abs/2303.12712>.
- Cash, P. & Maier, A. 2021 Understanding representation: contrasting gesture and sketching in design through dual-process theory. *Design Studies* 73, 100992; doi:[10.1016/j.des-tud.2021.100992](https://doi.org/10.1016/j.des-tud.2021.100992).
- Chen, L., Xia, D., Jiang, Z., Tan, X., Sun, L. & Zhang, L. 2024 A conceptual design method based on concept–knowledge theory and large language models. *Journal of Computing and Information Science in Engineering* 25 (2), 021001; doi:[10.1115/1.4066773](https://doi.org/10.1115/1.4066773).
- Cheng, H., Dai, Z., Liu, H., Yu, D., Lin, X. V., Liang, P., Tan, M.-W., Narang, S., Chung, H. W. & Le, Q. V. 2024. LLM factory: towards general and efficient methods for automatically constructing language models. Preprint. [arXiv:2407.04118](https://arxiv.org/abs/2407.04118).
- D’Souza, N. & Dastmalchi, M. R. 2017 “comfy” cars for the “awesomely humble”: exploring slang and jargons in a cross-cultural design process. In *Analysing Design Thinking: Studies of Cross-Cultural Co-Creation* (ed. E. Bohemia, T. Sung & M. Lemon), pp. 311–330. CRC Press.
- Daly, S. R., Christian, J. L., Yilmaz, S., Seifert, C. M. & Gonzalez, R. 2011 Assessing design heuristics for idea generation in an introductory engineering course. In *Proceedings of the American Society for Engineering Education Annual Conference (ASEE)*. <https://www.researchgate.net/publication/303856135>.

- Daly, S. R., Seifert, C. M., Yilmaz, S. & Gonzalez, R. 2016 Comparing ideation techniques for beginning designers: design heuristics, morphological analysis, and individual brainstorming. *Journal of Mechanical Design* **138** (10), 101108; doi:[10.1115/1.4034087](https://doi.org/10.1115/1.4034087).
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. 2019 Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the NAACL-HLT*, pp. 4171–4186; doi:[10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423).
- Doris, A. C., Grandi, D., Tomich, R., Alam, M. F., Ataci, M., Cheong, H. & Ahmed, F. 2024 Designqa: a multimodal benchmark for evaluating large language models' understanding of engineering documentation. *Journal of Computing and Information Science in Engineering* **25** (2), 021009; doi:[10.1115/1.4067333](https://doi.org/10.1115/1.4067333).
- Feng, W. B., Hélie, S. & Panchal, J. H. 2024 Using personas to increase the diversity of design concepts generated by large language models. *Proceedings of Design Computing and Cognition (DCC) 2024*, 71–88.
- Filippi, S. 2023 Measuring the impact of chatgpt on fostering concept generation in innovative product design. *Electronics* **12** (3535); doi:[10.3390/electronics12163535](https://doi.org/10.3390/electronics12163535).
- Fox, S. 2019 Addressing the influence of groupthink during ideation concerned with new applications of technology in society. *Technology in Society* **57**, 86–94; doi:[10.1016/j.techsoc.2018.12.009](https://doi.org/10.1016/j.techsoc.2018.12.009).
- Gero, J. S., Yu, R. & Wells, J. 2019 The effect of design education on creative design cognition of high school students. *International Journal of Design Creativity and Innovation* **7** (4), 196–212; doi:[10.1080/21650349.2019.1628664](https://doi.org/10.1080/21650349.2019.1628664).
- Girotra, K., Meincke, L., Terwiesch, C. & Ulrich, K. T. 2023 *Ideas are dime a dozen: Large language models for idea generation in innovation*. Working Paper, Mack Institute for Technological Innovation. <https://ssrn.com/abstract=4526071>.
- Goucher-Lambert, K. & Cagan, J. 2019 Crowdsourcing inspiration: using crowd generated inspirational stimuli to support designer ideation. *Design Studies* **61**, 1–29; doi:[10.1016/j.destud.2019.01.001](https://doi.org/10.1016/j.destud.2019.01.001).
- Grandi, D., Jain, Y. P., Groom, A., Cramer, B. & McComb, C. 2025 Evaluating large language models for material selection. *Journal of Computing and Information Science in Engineering* **25** (2), 021004; doi:[10.1115/1.4066730](https://doi.org/10.1115/1.4066730).
- Han, Y. & Moghaddam, M. 2024 Design knowledge as attention emphasize in large language model-based sentiment analysis. *Journal of Computing and Information Science in Engineering* **25** (2), 021007; doi:[10.1115/1.4067212](https://doi.org/10.1115/1.4067212).
- Hewstone, M., Rubin, M. & Willis, H. 2003 Intergroup bias. *Annual Review of Psychology* **53**, 575–604; doi:[10.1146/annurev.psych.53.100901.135109](https://doi.org/10.1146/annurev.psych.53.100901.135109).
- Hu, T. & Collier, N. 2024 Quantifying the persona effect in LLM simulations. Preprint. <https://arxiv.org/abs/2402.10811>.
- Janis, I. L. 1982 *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Houghton Mifflin.
- Jiang, C., Huang, R. & Shen, T. 2024 Generative ai enabled conceptualization: charting chatgpt's impacts on sustainable service design thinking with network-based cognitive maps. *Journal of Computing and Information Science in Engineering* **25** (2), 1–27; doi:[10.1115/1.4067211](https://doi.org/10.1115/1.4067211).
- Jonson, B. 2005 Design ideation: the conceptual sketch in the digital age. *Design Studies* **26** (6), 613–624; doi:[10.1016/j.destud.2005.03.001](https://doi.org/10.1016/j.destud.2005.03.001).
- Jutraz, A. & Zupancic, T. 2014 The role of architect in interdisciplinary collaborative design studios. *Intelligent Buildings International* **2014** (02), 34–42; doi:[10.15292/IU-CG.2014.02.034-042](https://doi.org/10.15292/IU-CG.2014.02.034-042).

- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., et al. 2023 Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and Individual Differences* **103**, 102274; doi:[10.1016/j.lindif.2023.102274](https://doi.org/10.1016/j.lindif.2023.102274).
- Kleinsmann, M., Deken, F., Dong, A. & Lauche, K. 2012 Development of design collaboration skills. *Journal of Engineering Design* **23** (7), 485–506; doi:[10.1080/09544828.2011.619499](https://doi.org/10.1080/09544828.2011.619499).
- Li, X., Sun, Y. & Sha, Z. 2025 Llm4cad: multi-modal large language models for three-dimensional computer-aided design generation. *Journal of Computing and Information Science in Engineering* **25** (2), 021005; doi:[10.1115/1.4067085](https://doi.org/10.1115/1.4067085).
- Li, Y., Ko, H. & Ameri, F. 2025 Integrating graph retrieval-augmented generation with large language models for supplier discovery. *Journal of Computing and Information Science in Engineering* **25** (2), 021010; doi:[10.1115/1.4067389](https://doi.org/10.1115/1.4067389).
- Liebrezn, M., Schleifer, R., Buadze, A., Bhugra, D. & Smith, A. 2023 Generating scholarly content with chatgpt: ethical challenges for medical publishing. *The Lancet Digital Health* **5**, e105–e106; doi:[10.1016/S2589-7500\(23\)00019-5](https://doi.org/10.1016/S2589-7500(23)00019-5).
- Luz de Araujo, P. H. & Roth, B. 2025 Helpful assistant or fruitful facilitator? Investigating how personas affect language model behavior. Preprint. <https://arxiv.org/abs/2407.02099>.
- Ma, K., Grandi, D., McComb, C. & Goucher-Lambert, K. 2023. Conceptual design generation using large language models. In *Proceedings of the ASME 2023 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC-CIE2023)*; doi:[10.1115/DETC2023-116838](https://doi.org/10.1115/DETC2023-116838).
- Ma, K., Grandi, D., McComb, C. & Goucher-Lambert, K. 2025 Do large language models produce diverse design concepts? A comparative study with human-crowdsourced solutions. *Journal of Computing and Information Science in Engineering* **25** (1), 024501; doi:[10.1115/1.4067332](https://doi.org/10.1115/1.4067332).
- McGee, R. W. 2023 *Who were the 10 best and 10 worst U.S. presidents? The opinion of ChatGPT (artificial intelligence)*. SSRN Electronic Journal; doi:[10.2139/ssrn.4367762](https://doi.org/10.2139/ssrn.4367762)
- Mikolov, T., Chen, K., Corrado, G. & Dean, J. 2013 Efficient estimation of word representations in vector space. Preprint. <https://arxiv.org/abs/1301.3781>.
- Monedero, J. 2000 Parametric design: a review and some experiences. *Automation in Construction* **9** (4), 369–377; doi:[10.1016/S0926-5805\(99\)00020-5](https://doi.org/10.1016/S0926-5805(99)00020-5).
- Naghavi, K. K., Wang, Z. & Xu, H. 2024 Reconstruction and generation of porous metamaterial units via variational graph autoencoder and large language model. *Journal of Computing and Information Science in Engineering* **25** (2), 021003; doi:[10.1115/1.4066095](https://doi.org/10.1115/1.4066095).
- Norheim, J., Rebentisch, E., Xiao, D. & de Weck, O. L. 2024 Challenges in applying large language models to requirements engineering tasks. *Design Science* **10**; doi:[10.1017/dsj.2024.8](https://doi.org/10.1017/dsj.2024.8).
- Osborn, A. F. 1957 *Applied Imagination: Principles and Procedures of Creative Thinking*. Charles Scribner's Sons.
- Plews-Ogan, M. L., Bell, T. D., Townsend, G., Canterbury, R. J. & Wilkes, D. S. 2020 Acting wisely: eliminating negative bias in medical education—Part 1: the fundamentals. *Academic Medicine* **95** (12 Supplement), S11–S15; doi:[10.1097/ACM.0000000000003699](https://doi.org/10.1097/ACM.0000000000003699).
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C. & Sutskever, I. 2023 Robust speech recognition via large-scale weak supervision. Preprint. <https://arxiv.org/abs/2303.09014>.
- Radford, A., Narasimhan, K., Salimans, T. & Sutskever, I. 2018 *Improving language understanding by generative pre-training*. OpenAI Technical Report. <https://cdn.openai.com/research-covers/language-understanding-pre-training.pdf>.

- openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.
- Rahutomo, F., Kitasuka, T. & Aritsugi, M. 2012 Semantic cosine similarity. In *Proceedings of the International Conference on Advanced Science and Technology (ICAST)*. https://www.researchgate.net/publication/262525676_Semantic_Cosine_Similarity.
- Sarica, S., Han, J. & Luo, J. 2023 Design representation as semantic networks. *Computers in Industry* **144**, 103791; doi:[10.1016/j.compind.2022.103791](https://doi.org/10.1016/j.compind.2022.103791).
- Thorp, H. H. 2023 ChatGPT is fun, but not an author. *Science* **379** (6630), 313; doi:[10.1126/science.adg7879](https://doi.org/10.1126/science.adg7879).
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E. & Lample, G. 2023 Llama: open and efficient foundation language models. Preprint. <https://arxiv.org/abs/2302.13971>.
- Tschimmel, K. 2012 Design thinking as an effective toolkit for innovation. In *Proceedings of the XXIII ISPIM Conference: Action for Innovation: Innovating from Experience*.
- Van Helden, G., Zandbergen, B. T. C., Specht, M. M. & Gill, E. K. A. 2023 Collaborative Learning in Engineering Design Education: A Systematic Literature Review. *IEEE Transactions on Education* **66** (5), 509–521; doi:[10.1109/TE.2023.3283609](https://doi.org/10.1109/TE.2023.3283609).
- Wang, B., Zuo, H., Cai, Z., Yin, Y., Childs, P., Sun, L. & Chen, L. 2023 A task-decomposed ai-aided approach for generative conceptual design. In *Proceedings of the ASME 2023 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC-CIE2023)*; doi:[10.1115/DETC2023-109087](https://doi.org/10.1115/DETC2023-109087)
- Wang, S., Moazeni, S. & Klabjan, D. 2025 A sequential optimal learning approach to automated prompt engineering in large language models. Preprint. [arXiv:2501.03508](https://arxiv.org/abs/2501.03508).
- Wu, C., Yin, S., Qi, W., Wang, X., et al. 2023 Visual ChatGPT: talking, drawing and editing with visual foundation models. Preprint. [http://arxiv.org/abs/2303.04671](https://arxiv.org/abs/2303.04671).
- Xie, H., Hoskins, D., Rowe, K. & Ju, F. 2025 Transformer-based offline printing strategy design for large format additive manufacturing. *Journal of Computing and Information Science in Engineering* **25** (2), 021011; doi:[10.1115/1.4067469](https://doi.org/10.1115/1.4067469).
- Yuksekgonul, M., Bianchi, F., Boen, J., Liu, S., Huang, Z., Guestrin, C. & Zou, J. 2024 TextGrad: automatic “differentiation” via text. Preprint. [arXiv:2406.07496](https://arxiv.org/abs/2406.07496).
- Zhang, C., Zhao, Y. F. & Haddad, R. E. 2024 Understanding the impact of applying large language models in engineering design education. *Journal of Computing and Information Science in Engineering* **25** (2), 021008; doi:[10.1115/1.4067318](https://doi.org/10.1115/1.4067318).
- Zheng, M., Pei, J., Logeswaran, L., Lee, M. & Jurgens, D. 2024 When “a helpful assistant” is not really helpful: personas in system prompts do not improve performances of large language models. Preprint. <https://arxiv.org/abs/2311.10054>.
- Zhu, Q. & Luo, J. 2023 Generative Transformers for Design Concept Generation. *ASME. J. Comput. Inf. Sci. Eng.* **23** (4), 041003; doi:[10.1115/1.4056220](https://doi.org/10.1115/1.4056220).