



Latest updates: <https://dl.acm.org/doi/10.1145/3643562.3672611>

RESEARCH-ARTICLE

Supermind Ideator: How Scaffolding Human-AI Collaboration Can Increase Creativity

JENNIFER L HEYMAN, Massachusetts Institute of Technology, Cambridge, MA, United States

STEVEN R RICK, Massachusetts Institute of Technology, Cambridge, MA, United States

GIANNI GIACOMELLI, Massachusetts Institute of Technology, Cambridge, MA, United States

HAORAN WEN, Massachusetts Institute of Technology, Cambridge, MA, United States

ROBERT J LAUBACHER, Massachusetts Institute of Technology, Cambridge, MA, United States

NANCY TAUBENSLAG, Massachusetts Institute of Technology, Cambridge, MA, United States

[View all](#)

Open Access Support provided by:

Polytechnique School

Mohammed VI Polytechnic University

Massachusetts Institute of Technology



PDF Download
3643562.3672611.pdf
29 December 2025
Total Citations: 15
Total Downloads: 3037

Published: 27 June 2024

[Citation in BibTeX format](#)

CI '24: Collective Intelligence Conference
June 27 - 28, 2024
MA, Boston, USA

Conference Sponsors:
SIGCHI

Supermind Ideator: How Scaffolding Human-AI Collaboration Can Increase Creativity

Jennifer L. Heyman*

Sloan School of Management,
Massachusetts Institute of Technology
USA

Steven R. Rick*

Sloan School of Management,
Massachusetts Institute of Technology
USA

Gianni Giacomelli*

Sloan School of Management,
Massachusetts Institute of Technology
USA

Haoran Wen

Sloan School of Management,
Massachusetts Institute of Technology
USA

Robert J. Laubacher

Sloan School of Management,
Massachusetts Institute of Technology
USA

Nancy Taubenslag

Sloan School of Management,
Massachusetts Institute of Technology
USA

Max Sina Knicker

Econophysix and Complex Systems,
Ecole Polytechnique
France

Younes Jeddi

School of Collective Intelligence,
Mohammed VI Polytechnic University
Morocco

Pranav Ragupathy

Sloan School of Management,
Massachusetts Institute of Technology
USA

Jared Curhan

Sloan School of Management,
Massachusetts Institute of Technology
USA

Thomas W. Malone

Sloan School of Management,
Massachusetts Institute of Technology
USA

ABSTRACT

Previous efforts to support creative problem-solving have included (a) techniques such as brainstorming and design thinking to stimulate creative ideas, and (b) software tools to record and share these ideas. Now, generative AI technologies can suggest new ideas that might never have occurred to the users, and users can then select from these ideas or use them to stimulate even more ideas. To explore these possibilities, we developed a system called Supermind Ideator that uses a large language model (LLM) and adds prompts, fine tuning, and a specialized user interface in order to help users reformulate their problem statements and generate possible solutions. This provides scaffolding to guide users through a set of creative problem-solving techniques, including some techniques specifically intended to help generate innovative ideas about designing groups of people and/or computers (“superminds”). In an experimental study, we found that people using Supermind Ideator generated significantly more innovative ideas than those generated by people using ChatGPT or people working alone. Thus our results suggest that the benefits of using LLMs for creative problem-solving can be substantially enhanced by scaffolding designed specifically for this purpose.

*Authors Heyman, Rick, and Giacomelli contributed equally.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CI '24, June 27–28, 2024, Boston, MA, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0554-0/24/06
<https://doi.org/10.1145/3643562.3672611>

CCS CONCEPTS

- Computing methodologies → Artificial intelligence; • Applied computing; • Human-centered computing → Human computer interaction (HCI); HCI design and evaluation methods; Interaction paradigms;

KEYWORDS

Creativity, Innovation, Collective Intelligence, Generative AI, Scaffolding, Large Language Models

ACM Reference Format:

Jennifer L. Heyman, Steven R. Rick, Gianni Giacomelli, Haoran Wen, Robert J. Laubacher, Nancy Taubenslag, Max Sina Knicker, Younes Jeddi, Pranav Ragupathy, Jared Curhan, and Thomas W. Malone. 2024. Supermind Ideator: How Scaffolding Human-AI Collaboration Can Increase Creativity. In *Collective Intelligence Conference (CI '24), June 27–28, 2024, Boston, MA, USA*. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3643562.3672611>

1 INTRODUCTION

Creative problem-solving is critical to success in many kinds of human activity, from architecture, engineering, and public policy to art, entrepreneurship, and designing the human groups that perform all these activities [9, 14, 19, 41, 50]. It is, therefore, not surprising that many techniques to improve creative problem-solving have been proposed over the years, including brainstorming, design thinking, mind-mapping, crowdsourcing, and many others [18, 23, 49].

In this paper, we investigate the potential of a new kind of tool—generative AI—for supporting creative problem-solving. In particular, we focus on how large language models (LLMs)—such as the GPT (Generative Pre-trained Transformer) family of LLMs [5]—can take natural language descriptions of a problem as input and

produce as output natural language ideas about how to reframe or solve the problem.

Interestingly, even though one widely discussed limitation of today's LLMs is that they sometimes produce incorrect or irrelevant outputs, this limitation is usually not a problem when the generative AI system is used to *augment* human creativity instead of *replacing* it. In this case, human users can often easily decide which of the outputs from the system are useful enough to consider further and which aren't. And even ideas that may at first seem irrelevant can sometimes trigger further useful ideas for human users. In fact, trying to make connections between a problem and seemingly unrelated ideas is one simple technique for triggering creative ideas [35].

However, it is still unclear how best to support the use of generative AI technologies for creative problem-solving and whether their usefulness can be substantially improved by scaffolding to help users carry out structured processes. To study this question, we developed the Supermind Ideator, a system that uses an LLM with specialized prompts, fine-tuning, and a user interface to help users reflect upon their problems and generate possible solutions. The system does this using a set of conceptual *moves*—techniques that humans can use to trigger creative ideas. By sequencing these moves in different orders and combinations, users can explore many different ideas for a given problem.

Most of the techniques we currently use in the Supermind Ideator are based on the "Supermind Design" methodology [17]. Some of these techniques, such as looking at sub-parts or analogies, can be helpful for addressing any problem. Other techniques are specifically intended to help generate innovative ideas about how to design *superminds*, defined as *groups of individuals acting together in ways that seem intelligent* [39]. For instance, one such supermind design technique encourages users to consider how a problem could be solved with different kinds of groups, such as hierarchies, democracies, markets, or communities.

In other words, "superminds" is a short way of saying "collectively intelligent systems." Thus the Supermind Ideator illustrates how a collectively intelligent system composed of a person and a computer can design other collectively intelligent systems. The rest of this paper is about the implementation of the Supermind Ideator, how people used it to help them generate creative ideas, and how that process compared with two other standard practices: (a) humans working alone and (b) humans using ChatGPT.

2 RELATED WORK

Prior work has extensively discussed creative problem-solving approaches that use techniques from design thinking to collective intelligence. These approaches predominantly focus on groups using organizational and methodological approaches to address issues such as design fixation, knowledge curation, and creative inspiration.

2.1 Facilitated Idea Generation

From the early days of design being applied in the scientific arena [11] to the modern-day frameworks for using Design Thinking methodologies [12], creative problem-solving techniques have evolved significantly. For example, the mix of deep problem understanding

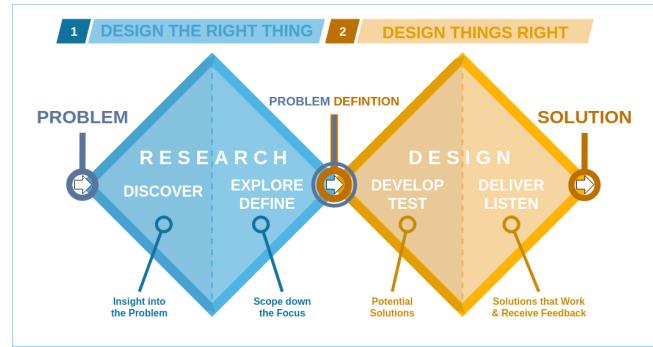


Figure 1: The Double Diamond method highlights two components of creative problem-solving (understanding the problem and developing a solution) and showcases the four steps where one first discovers and defines the problem before then developing and delivering the solution. Source: Wikipedia - Double Diamond (design process model)

and iterative solution generation, most notably combined in the Double Diamond method [1], enables a rigorous and empathetic approach to integrate the needs of people, the possibilities of technology, and the requirements for business success. In the first of the two diamonds, practitioners (a) *diverge* by considering different ways to frame the problem and then (b) *converge* to narrow down to a useful problem definition. Then, in the second diamond, they (a) *diverge* by considering different potential solutions to the problem and then (b) *converge* on a few of the best solutions, as shown in Fig 1.

Other prior work has studied the phenomenon of creative ideation through sociocultural lenses [14, 18, 19, 37, 46], digging into specific topics and issues such as design fixation [9, 36, 52, 59], bias [41], inspiration [15, 55], and innovation [24, 46]. According to Gabora [19], for instance, ideas emerge, evolve, and manifest as creative products through the formation of combinations and reorganization of existing ideas. The process can be modeled by constraint-based iteration to transform ideas into tangible solutions.

However, idea-generation activities can be long and laborious, and mental tendencies, such as functional fixedness, often lead to design fixation [26]. Design fixation is a cognitive bias that presents as an over-fitting of design space to pre-existing knowledge and experiences of the designer [9]. It hinders wide-reaching ideas and overly constrains the domain of idea generation, functionally falling into a local maximum and missing the wider absolute maximum available. Studies have attempted to develop prevention methods such as concept mapping, remote association, facilitated design thinking, and exposure to other new and unrelated content [24, 52, 59]. More recent work has examined the potential for collective intelligence approaches such as crowdsourcing idea generation, crowd-based ratings, and other forms of computer-supported cooperative work to help alleviate design fixation and access diverse types of knowledge [22, 28, 31, 34, 38, 60]. All of these approaches, however, still largely rely on human effort and access to knowledge resources that do not easily scale.

2.2 Generative AI

The advent of Transformer architectures in language modeling has led to a meteoric rise in the use of LLMs [57]. New systems like OpenAI's GPT systems have facilitated general access to inferencing models that rapidly produce large-scale human-quality text output from a small input [5]. These models can be tuned and directed with either (a) simple instructions (*zero-shot learning*), (b) small numbers of example input and output pairs (*few-shot learning*), or (c) large numbers of example inputs and outputs (*fine-tuning*). With these elements, LLMs are capable of very wide-ranging idea generation [54], and researchers are beginning to understand how to improve the quality of output from LLMs to better match the design space constraints [61]. Prior research suggests that generating more ideas can lead to greater creativity [45] so it seems plausible that using generative AI would help humans generate more ideas and in turn lead to greater creativity.

To our knowledge, little systematic work has explicitly explored how to combine the expertise of humans with the capabilities of generative AI technologies. However, given that current generative AI systems by themselves are often not capable of producing uniformly high-quality ideas, a very promising possibility is to let generative AI systems rapidly generate ideas for humans to consider. Generative technologies are already capable of providing usefully unexpected inputs and stimuli for human designers. And this can increase the range of possibilities to be considered by humans. Thoughtfully designing the way a system guides and facilitates this process can thus improve the probability that designers will find better solutions faster than they would have otherwise.

Two core facets underlie this concept: *reflection* and *inspiration*.

2.2.1 Reflection. During creative problem-solving, it is often useful to engage in a reflective feedback loop [10, 58]. While random feedback can be much better than no feedback, an approach that facilitates reflection can prevent early fixation on solutions and help explore design space constraints from both the problem and solution domains [42].

A critical concern with generative AI, and LLMs in particular, is that end users need to be able to articulate their goals in a way that leads the model to produce the desired output [53]. In the same way that Hutchins, Hollan, and Norman [25] discuss minimizing the distance between a user's thoughts and the user interface they are manipulating, it is important that generative systems provide indications to the end user of what the system can do and how to do it.

We extend this perspective by observing that much of the work done by design thinking and innovation facilitators is less an act of solution generation and curation and more of prompting insight through introspective questions and reflective actions. With this in mind, we suggest a useful characteristic for any innovation support system - it should promote reflection by humans rather than impose unnecessary constraints on their thinking. For instance, considering system outputs as "ideas" rather than "answers" and positioning the system as a tool to *support* human cognition rather than *replace* it may be key. And it seems especially important to emphasize to the human users that they are responsible for deciding what to do with the system's outputs. They can ignore the system outputs, use them

without change, or let them stimulate further creative thinking by the users.

2.2.2 Inspiration. The second critical quality for any innovation support system is to promote wide-reaching exploration of the available design space. This is often accomplished through facilitated activities that promote divergent rather than convergent thinking, such as the alternative uses test [20]. Many divergent thinking activities put the onus of creating new uses on the human engaging in the activity [7]. Fortunately, it is increasingly being found that generative technologies like GPT can do quite well to replicate human idea generation within such divergent thinking tasks [54]. This does not replace human ideation. Instead, it presents a chance to augment human idea generation and avoid design fixation by exposing humans to novel and diverse stimuli that stimulate new connections [24, 52, 59].

A complex topic concerning LLMs today is their safety and accuracy with generated content [3, 51]. Transformer-based systems are capable of exceptionally human-like language generation, but they do not always produce factual content nor can they reliably perform reasoning tasks in their current state [56]. These "hallucinations" by the machine can sound as if they carry a sense of authority and accuracy even when they are simply false combinations of words that statistically sound appropriate given prior text and the underlying model of the LLM [21].

We see this (and much of the rest of the content that LLMs can create) more as an opportunity and feature rather than a bug. Outlandish and unfeasible ideas are a natural fit for idea-generation practices. By interrogating the extremes of a design space, it is possible to construct a better understanding of what is actually possible and appropriate. By framing all content as ideas (including both good and bad ideas), we present the designer with a more far-reaching design space exploration. Such support for divergent thinking is essential for the success of any creative problem-solving support system.

3 THE SUPERMIND DESIGN METHODOLOGY

The Supermind Ideator is based on the Supermind Design methodology, which includes a set of conceptual moves that people can use to spur their creativity, especially about how to design collectively intelligent groups [17]. These moves have been used successfully in multiple settings [30, 32, 33].

The methodology includes the following *basic design moves* that are based on general techniques that can be used for any kind of creative problem-solving [2, 4, 13, 29, 40, 44]:

- *Zoom In - Parts*: What are the parts of this problem?
- *Zoom In - Types*: What are the types of this problem?
- *Zoom Out - Parts*: What is this problem a part of?
- *Zoom Out - Types*: What is this problem a type of?
- *Analogize*: What are analogies for this problem?

The methodology also includes the following *supermind design moves* that are specifically for generating ideas about how to design superminds (i.e., collectively intelligent groups) [17, 39]:

- *Groupify*: How can different kinds of groups help solve the problem? Possibilities include:

- *Hierarchy* - where group decisions are made by delegating them to individuals in the group
- *Democracy* - where group decisions are made by voting
- *Market* - where group decisions are the combination of all the pairwise agreements between individual buyers and sellers
- *Community* - where group decisions are made by informal consensus based on shared norms and reputations
- *Ecosystem* - where group decisions are made by whoever has the most power and by survival of the fittest
- *Cognify*: How can different cognitive processes help solve the problem? Possibilities include:
 - *Create* - How can groups create new information collectively?
 - *Decide* - How can groups make choices?
 - *Sense* - How can groups collect and interpret information from the environment?
 - *Remember* - How can groups recall information from the past?
 - *Learn* - How can groups improve their performance with experience?
- *Technify*: How can technologies be used to help solve the problem?

Finally, the methodology includes three *experimental moves* which have not, to our knowledge, been previously used as part of systematic ideation exercises but which appear to take advantage of LLM capabilities and are incorporated in the Supermind Ideator:

- *Reflect* - What is missing from the current problem statement?
- *Reformulate* - How could the problem be reformulated?
- *Case examples* - How does the problem relate to case examples of real groups?

4 THE SUPERMIND IDEATOR SOFTWARE

The Supermind Ideator¹ is a web application designed to guide users into a focused state of idea generation and reflection, and as such is intentionally kept minimalist, as shown in Fig 2. The left side consists of a 'Generate Panel' that drives all of the idea creation on the right side. After users type in their problem, they are given three options for idea creation: *Explore Problem*, *Explore Solutions*, and *More Choices*. These options are meant to provide scaffolding for novice users who may not know where to begin, as well as to support more advanced users who already know what they want to do next. The first two options, *Explore Problems* and *Explore Solutions*, comprise what we have called "move sets," or groups of moves that focus on a specific aspect of the idea generation and refinement process.

The *Explore Problem* move set supports the problem definition part of the double diamond approach using the *basic design moves* and the *experimental moves* from the Supermind Design methodology. In this way, it helps users reflect on how they can generalize and specialize the various parts and types of their problem, consider relevant analogies to their problem, and identify potentially missing aspects of their problem statement.

¹<https://ideator.mit.edu/>

The *Explore Solutions* move set supports the solution generation part of the double diamond approach using the *supermind design moves* from the Supermind Design methodology. For example, it helps users consider how different kinds of groups (such as markets, communities, and democracies) could help solve their problem. It also helps users think about how innovative ways of performing different cognitive processes (such as creating, deciding, and sensing) or using various kinds of technologies could help solve their problem.

These "move sets" take the user input and then execute several moves in order to create several ideas, one output for each move.

More Choices allows a user to select any individual move(s) they want. The *More Choices* option also exposes a more advanced parameter that the LLM API calls "temperature" - a measure of the amount of randomness used to generate output. Lower temperature leads to less random (more conventional) outputs, and higher temperature leads to more random (more potentially creative) outputs. To avoid confusion, we name this "Creativity" and provide three choices: Low, Medium, and High, corresponding to temperatures of 0.7, 1.0, and 1.3, respectively.

In each of the above cases, the Ideator system generates one or more ideas using the Supermind Design moves. After ideas are created and appear on the right, users can rate each of these ideas with a Thumbs Up or Thumbs Down button, and they can Bookmark ideas they particularly like to save these ideas in their personal collection.

4.1 Implementation

The Ideator system includes multiple layers that work together to provide structure and guidance to an otherwise open-ended problem. These layers (shown in Fig 3) include the *User Interface*, the *API*, and the *LLM*.

To increase the ease of maintaining and extending the system, we built an API (Application Programming Interface) to separate the client-side user interface logic from server-side processing logic.

We expect this API-based approach to also be valuable in making the Supermind Ideator moves available to others who might want to build upon or extend our methodology beyond the initial web interface we have provided. The API simplifies each move into a simple set of text strings with most of the formatting and interaction with the LLM abstracted away to optional parameters.

The API finally allows for the logic of each move to be isolated and version-controlled to support experimentation and evolution. As we improve our understanding of prompting techniques and as the capabilities of LLMs expand and change, we have infrastructure in place to support upgrading and adapting moves.

4.2 Prompting

To tell the LLM how to respond to a given problem using a given move, we use two kinds of *prompts*: "zero-shot" and "few shot." In a zero-shot prompt for a move, the API simply gives English instructions to the LLM for how to apply the move. For example, the zero-shot prompt for the *Groupify - Democracy* move is: "Democracies are systems where group decisions are made by voting. How can we use democracies to {problem}? To solve this problem, ". The API inserts the user's problem description in place of "{problem}",

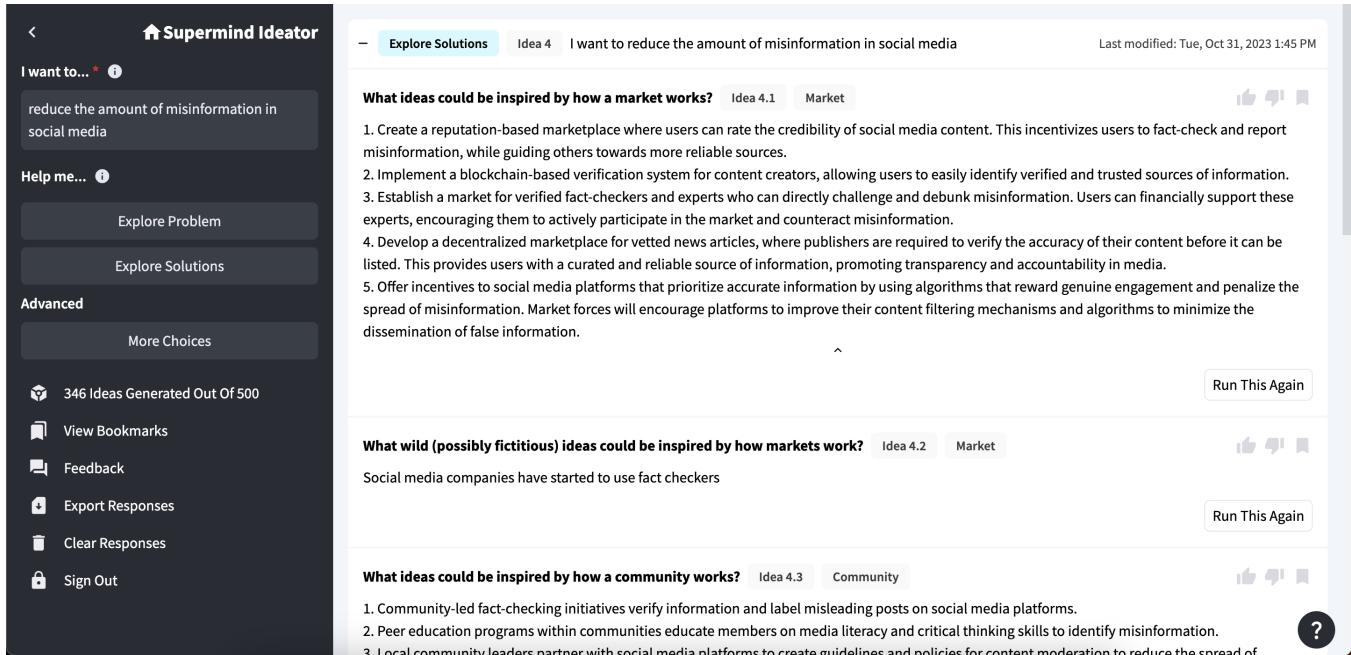


Figure 2: The Supermind Ideator Interface. The left side contains the Generate Panel where users input their problem and select Moves to run. The right side contains ideas generated by the system.

and the LLM then generates text that would be likely to follow this prompt.

In a few-shot prompt for a move, the API gives the LLM several examples of problems and desirable responses to those problems, followed by the user's problem. Then the LLM generates a response to the user's problem following the same pattern of problem/response that is illustrated by the examples.

For example, with Zoom In - Parts, we provide multiple examples of how to break down a problem into its parts. One such example is:

- What are the parts of my problem? Problem: sell
- Identify potential customers
- Identify potential customer needs
- Inform potential customers
- Obtain order
- Deliver product or service
- Receive payment

4.3 Fine-Tuning

While the general knowledge embedded within the LLM base model can produce interesting and fitting output with relatively limited guidance (zero-shot and few-shot prompts), it seemed apt to augment this base knowledge with specialized information so that more precise and useful ideas could also be generated. To accomplish this, we curated approximately 1600 examples of real-world organizational practices that solved problems using various innovative approaches. We organized this case study corpus with input labels that connected the problem and case study with an associated Cognify or Groupify move.

This enabled us to fine-tune a version of the LLM so a new model could map between input language, output prediction, and Supermind move. This new fine-tuned model is much less constrained than our zero-shot and few-shot moves. While these earlier moves answer a question or problem with another question or idea based on the base LLM and our prompting, the fine-tuned approach changes the LLM so that output more tightly reflects the input training data. As our training data was a corpus of case studies, the fine-tuned model produces output more akin to those case studies. This produces an interesting challenge and opportunity as the LLM tends to produce fictional content in a coherent and sensible manner, reading as if it is stating facts.

For example, when prompted with the problem: "I want to improve the way companies retrain employees whose jobs have been replaced by automation," we get examples like the following:

- Groupify - Market: "Swiss logistics company Travida uses a collective intelligence method to help with employee retraining in the face of automation replacement. Its platform, called Morpheus, allows for crowdsourcing of training needs, as well as for matching of people with training openings"
- Cognify - Remember: "Microsoft has been experimenting with virtual retraining environments that mix real and virtual reality to allow workers to experience new scenarios. The idea is that learning should be less dependent on physical location, and could be reimaged as a continuous, fluid process that doesn't end with graduation."

Even though both of these examples sound factual, they are actually "hallucinations," fictional combinations of information from different case studies in our corpus, the base LLM, and other predictions made by the transformer's neural network.

If our goal were to help users retrieve factual content, this would be a problem. But as noted above, since our goal is to help users come up with creative ideas, we think these fictitious hallucinations are more likely to be a feature than a bug. Creativity researchers have found, for example, that exposure to other creative ideas can often be useful triggers for stimulating people to have creative ideas they would never have had otherwise [16]. To avoid misconceptions on the part of our users, however, we label these outputs as "wild (possibly fictitious) ideas" which can be seen in Fig 2.

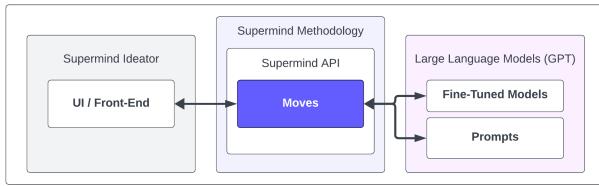


Figure 3: The Layers of the Supermind Ideator System. These consist of the User Interface, the API, and the LLM being used. The interface is on top (left) and is based on a React framework. The API is in the middle and connects the interface to the LLM through the software that implements our Moves. The LLM is at the bottom (right) and includes special prompts and specially fine-tuned versions of the LLM.

5 EXPERIMENTAL STUDY

In order to evaluate the system, we conducted an experimental study of how Ideator can facilitate humans in creative problem solving tasks. At the time this study took place, Ideator was configured to make use of OpenAI's GPT-3.5 Turbo LLM.

5.1 Methods

5.1.1 Procedure. To examine how Ideator facilitates humans in creative problem solving tasks, we employed a two phase (Idea Generation and Idea Evaluation) study. In the Idea Generation phase, we employed a 2 (problem statement) by 3 (group configuration) mixed-factorial design with problem statement being randomized within-person and group configuration being randomized between-person. Participants completed two rounds of a creative problem solving task in which they were asked to generate as many ideas as they could to solve the counterbalanced problem statements "How can I discern fake news from real news" and "How can I achieve better work-life balance". Participants had no time limit and were free to spend as much or as little time as needed to complete the task successfully.

5.1.2 Study Phases. The study involved three phases: Baseline, Idea Generation, and Idea Evaluation.

Baseline. In the first round of this task, all participants were asked to generate ideas without any help from external sources (i.e., not using Google, ChatGPT, etc). This was used as a baseline level of performance on the creative problem solving task ($N_{Ideas} = 611$).

Idea Generation. Participants were then randomly assigned to one of three group-configuration conditions and asked to complete the creative problem solving task again with the other problem

statement. Specifically, participants were assigned to one of three conditions:

- Human + Ideator condition, in which they were asked to generate as many ideas as possible with the help of Ideator ($N_{Participant} = 51, N_{Ideas} = 200$)
- Human + ChatGPT condition, in which they were asked to generate as many ideas as possible with the help of ChatGPT ($N_{Participant} = 52, N_{Ideas} = 231$)
- Human Only condition, in which they were asked to generate as many ideas as possible without any help from external sources ($N_{Participant} = 49, N_{Ideas} = 172$).

Idea Evaluation. The submitted ideas were finally evaluated by a separate group of participants who rated each solution based on its level of innovativeness (i.e., creativity and usefulness) using a 1 (not at all innovative) to 5 (extremely innovative) scale. Each evaluator viewed a random subset of 100 ideas and each was rated by at least 17 evaluators. The intra-class correlation coefficient was computed to assess the agreement between evaluators who rated the innovativeness of the same ideas. There was a good reliability between the evaluators using the one-way random effects model and "absolute agreement" using R's ICC function ($\kappa = 0.81, p < .0001$).

5.1.3 Participants. Participants were recruited from the online crowdsourcing site Prolific. All participants were located in the United States and were fluent in English. We recruited a total of 152 idea generators ($M_{Age} = 37.19$ years, $SD_{Age} = 12.93$ years, 53.3% Females, 45.4% Males, 1.3% Other Sex, 9.9% Asian, 10.5% Black, 63.8% White, 10.5% Mixed, 5.3% Other Ethnicity) and 505 idea evaluators ($M_{Age} = 40.83$ years, $SD_{Age} = 13.86$ years, 62.8% Females, 35.4% Males, 1.8% Other Sex, 7.3% Asian, 12.5% Black, 65.5% White, 7.7% Mixed, 7.0% Other Ethnicity).

5.2 Results

5.2.1 Does Innovativeness Vary by Condition? To examine whether innovativeness ratings vary by condition, we conducted a one-way analysis of variance. However, our data did not meet the homogeneity of variances assumption. Therefore, we conducted a Welch's one-way ANOVA and a Games-Howell post-hoc test.

We found that there was a significant difference between the Human + Ideator, Human + ChatGPT, and Human Only condition in terms of their innovativeness ratings ($F(2, 8058.15) = 106.48, p < .0001$). Specifically, the Human + Ideator ideas were rated as being the most innovative ($M = 3.21, SD = 1.26$), followed by Human + ChatGPT ($M = 2.96, SD = 1.24$) and Human Only ($M = 2.81, SD = 1.27$). All comparisons were significant (all $p < .0001$; see Figure 4).

5.2.2 Does Baseline Performance Moderate the Relationship Between Condition and Innovativeness? To examine whether baseline performance on the idea generation task moderated the relationship between condition and innovativeness, we added baseline performance as a covariate in the above ANOVA. Additionally, we conducted a linear regression with innovativeness being regressed on both condition and baseline performance.

We found that performance on the baseline creative problem solving task significantly predicted better performance on the experimental task for the Human Only condition ($b = 0.63, z = 3.86, p$

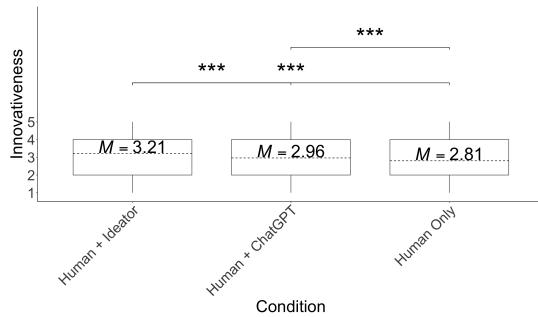


Figure 4: Comparison between the different experimental conditions in terms of Innovativeness ratings. * $p < .0001$**

= .0002). However, performance on the baseline creative problem solving task did not significantly predict better performance on the experimental task for those in the Human + Ideator ($b = 0.20$, $z = 1.43$, $p = .16$) or Human + ChatGPT conditions ($b = 0.31$, $z = 1.81$, $p = .07$).

In other words, subjects in the Human Only condition who perform better on the baseline task also perform better on the experimental task (see Figure 5).

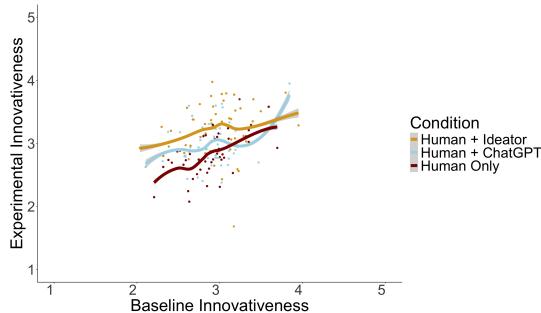


Figure 5: Association between the innovativeness of a subject's performance on the Baseline task and the Experimental task. Each point represents one subject, and the colored lines are best-fit lines for each of the three conditions.

5.2.3 Number of conversational turns. By tracking individual usage of the two treatment systems (Ideator and ChatGPT), we were able to compare how many times the users interacted with the system, something we equate to 'conversational turns' with the machine. In the Human + ChatGPT condition, we count one turn for each time a user sent a message and received a response. In the Human + Ideator condition, we count one turn for each time a user submitted a problem statement (to Explore Problem, Explore Solution, or other individual moves) and received back one or more ideas.

A Wilcoxon rank-sum test indicated that participants in the Human + Ideator condition ($M = 3.11$, $Mdn = 2$) went through a significantly greater number of conversational turns compared to the Human + ChatGPT condition ($M = 1.54$, $Mdn = 1$; $W = 12278.50$, $p < .0001$, effect size $r = 0.42$). In other words, those using Ideator tended to take approximately 3 conversational turns before

finishing the creative problem solving task, whereas those using ChatGPT tended to take less than 2 conversational turns before finishing the task (see Figure 6).

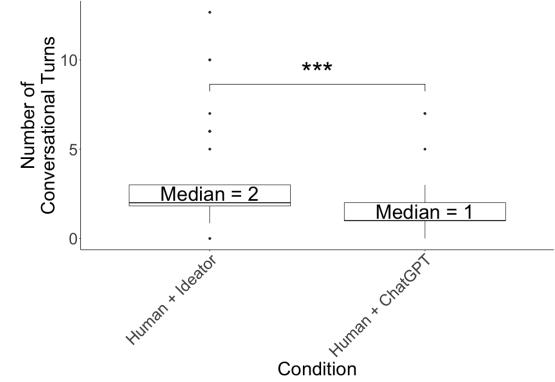


Figure 6: Comparison between the Human + ChatGPT and Human + Ideator conditions in terms of the number of conversational turns idea generators typed in.

Since one problem statement with Ideator generated more ideas than one problem statement with ChatGPT, this means that the Ideator users saw significantly more software-generated ideas than the users of ChatGPT ($M_{Ideator} = 17.00$, $Mdn_{Ideator} = 13$; $M_{Chat} = 1.54$, $Mdn_{Chat} = 1$; $W = 1329.00$, $p < .0001$, effect size $r = 0.83$ - see Figure 7).

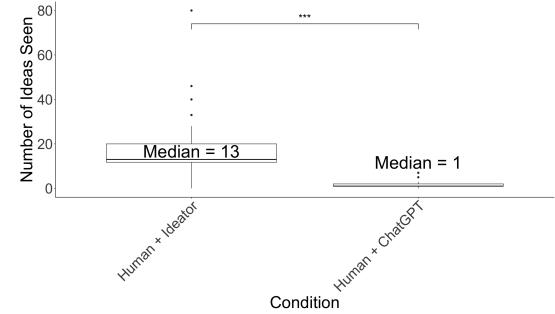


Figure 7: Comparison between the Human + ChatGPT and Human + Ideator conditions in terms of the number of ideas seen by each idea generator.

Additionally, we examined whether there was a significant difference in the number of ideas submitted between the Human + Ideator, Human + Chat, and Human Only conditions. A Kruskall-Wallis test indicated that there was no significant difference in the number of ideas submitted between those in the Human + Ideator ($M = 3.92$, $SD = 2.64$), Human + ChatGPT ($M = 4.44$, $SD = 2.39$), or Human Only ($M = 3.51$, $SD = 1.49$) conditions ($H = 3.59$, $p = .17$ - see Figure 8).

Overall, this means that the users in the Human + Ideator condition had more ideas to choose from in selecting the ideas they submitted, and this may help explain why their submitted ideas were rated as higher quality.

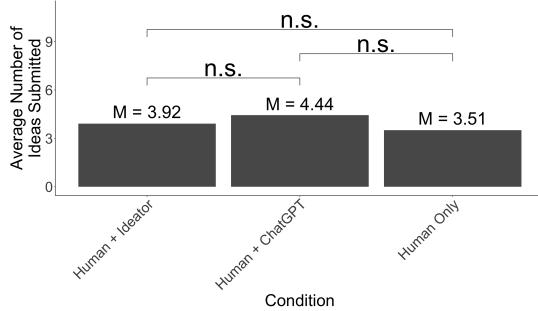


Figure 8: Comparison between the experimental conditions in terms of the average number of ideas submitted by idea generators.

5.2.4 Copy and Paste. To examine whether the extent to which participants copied and pasted responses directly from the two treatment systems (Ideator and ChatGPT) differed between the Human + Ideator and Human + ChatGPT conditions, we conducted a chi-square test of independence. We found that those in the Human + Ideator condition tended to copy and paste their answers from the system significantly more frequently than those in the Human + ChatGPT condition. Specifically, 30% of solutions submitted from the Human + Ideator condition were copy and pasted directly from the system, whereas 18.18% of ideas generated from the Human + ChatGPT were copy and pasted ($\chi^2 (1, 431) = 7.65, p = .006$ - see Figure 9). This suggests that the human users found ideas generated by Ideator more directly usable than those generated by ChatGPT.

Furthermore, we conducted an independent sample t-test to examine whether the innovativeness ratings differed between ideas that were copy and pasted and ideas that were not. We found that ideas that were copy and pasted tended to be rated as being more innovative, although this effect was stronger for those in the Human + Ideator condition ($t = 10.94, p < .0001$) than those in the Human + ChatGPT condition ($t = 2.73, p = .006$).

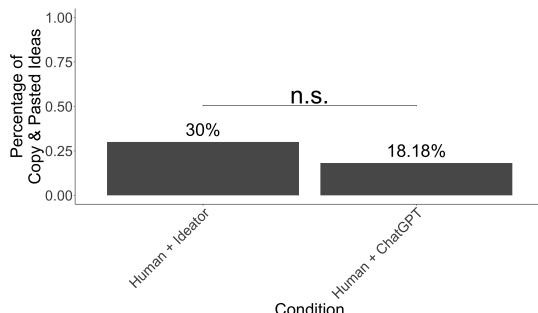


Figure 9: Comparison between the Human + ChatGPT and Human + Ideator conditions in terms of the percentage of ideas copy and pasted directly from the system.

5.2.5 Moves. Across all 30 moves that Ideator can present to help users process their problem statements, the results of only 13 moves were copy-pasted directly from the system and submitted as ideas.

Interestingly, the most and least innovative of these moves differed depending on the problem statement that participants were exploring. Specifically, when exploring the work-life balance problem statement, moves generated from Cognify - Create (Zero-shot) were rated as being the most innovative ($M = 3.80, Mdn = 4$), whereas moves generated from Zoom Out - Parts were rated as being the least innovative ($M = 2.63, Mdn = 3$). In contrast, when exploring the fake news problem statement, the moves generated by Technify (Zero-shot) were rated as being the most innovative ($M = 3.89, Mdn = 4$) whereas the moves generated by Analogize were rated as being the least innovative ($M = 2.86, Mdn = 3$).

Diving into this more deeply, we found that innovativeness of submitted ideas generated by the various implementations of moves (few-shot, fine-tune, and zero-shot) differed depending on which problem statement participants were exploring. Specifically, when exploring the work-life balance problem statement, there was no significant difference in innovativeness ratings of moves generated from few-shot ($M = 3.56, Mdn = 4$), fine-tune ($M = 3.16, Mdn = 3$), or zero-shot ($M = 3.54, Mdn = 4$; $F(2, 104.67) = 2.56, p = .08$). However, when exploring the fake-news problem statement, moves generated from zero-shot were rated as being significantly more innovative than moves generated from few-shot ($M_{zero-shot} = 3.89, M_{few-shot} = 3.57, t = 3.23, p = .001$). No ideas were submitted from fine-tuned moves for those who saw the fake-news problem statement.

Together, these results suggest that the most helpful moves might differ depending on the type of problem users are exploring when using Ideator. We discuss this in greater detail in the discussion section.

In addition, the frequency with which generated ideas were submitted as final ideas differed by the type of move implementation ($\chi^2 (2, 51) = 35.41, p < .0001$; see Figure 10). The most common was few-shot ($N = 36$), followed by zero-shot ($N = 13$), followed by fine-tune ($N = 2$). This suggests that the users did not highly value the results from fine-tune moves. And since creating the training data for the fine-tune moves requires much more effort than creating the zero-shot and few-shot moves, this suggests that, at least for the fine-tune examples we used, the additional effort may not have been worth the cost.

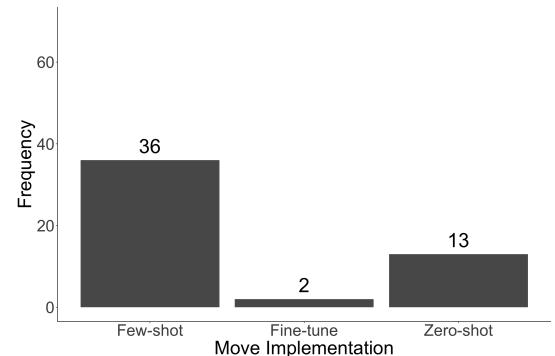


Figure 10: Frequency of move implementation in submitted ideas

6 DISCUSSION

As generative AI systems (and LLMs in particular) continue to be developed and deployed in real world applications, our work suggests several implications for researchers who study these technologies and practitioners who apply them. In particular, we show how specialized front ends that support systematic processes for specific problem types can scaffold human-AI interactions in useful ways.

6.1 Specialized front ends

While chat-based LLMs are usually very easy to use and can be used for a wide range of applications, this work demonstrates that there are also situations where specialized front ends can be valuable. As our results show, human users of Ideator were better able to generate innovative solutions to complex problems than human users of ChatGPT.

We believe this is partly because most users are not expert prompt engineers, and they may not be very good at using a chat-based interface to carry out a structured process for a specific kind of task. But as the Supermind Ideator illustrates for the task of creative problem-solving, a specialized front end can include (a) carefully selected prompts that are known to achieve good results for this task and (b) customized UIs that greatly simplify using these prompts in a structured process with inputs specific to a user's problem.

Some commercial products are now incorporating capabilities like these (e.g., OpenAI's custom GPTs and Assistants), but Supermind Ideator has included them for some time [48], and the current work experimentally demonstrates their value in the domain of creative problem-solving.

6.2 Systematic processes

For tasks with codified methodologies that humans use, there is value in guiding users through a process. Unlike a chat-based interface that tries to gratify the user at each step of the chat, a specialized front-end can act more like a human facilitator who guides the user through the activities specified in the methodology. For example, in the creative problem-solving domain, human facilitators often take participants through a long process to break them out of their own fixations and find new ways of understanding the problem before exploring possible solutions. The Supermind Ideator supports this process without human facilitators by guiding users to explore the problem space first followed by the solution space. Indeed, we found that users of Ideator tended to engage with the system more than users of ChatGPT.

Importantly, however, approximately one-third of Ideator users tended to take results from Ideator verbatim, rather than using the results as a jumping-off point to further revise and build upon. We believe this role for humans as "curators" of ideas generated by AI is an important one, and our results showed that this was associated with high ratings of innovativeness.

In fact, this may be one reason why the users who didn't perform as well on the Baseline task appeared to benefit more from using Ideator (see Figure 5). Even if these users didn't have very many innovative ideas themselves, they could still recognize and submit innovative ideas generated by Ideator. This effect of generative AI

benefiting lower-skilled workers more than higher-skilled ones has been found by a number of other researchers [6, 8, 43].

In the future, however, we also hope to encourage Ideator users to reflect and iterate upon the Ideator results which, ultimately, may contribute to even better ideas being formed.

6.3 Problem Types

There is a near-infinite range of complex problems in the world, from personal to socio-political. And different types of problems can often benefit from specialized knowledge or approaches. For example, the Ideator includes the three supermind design moves, which are especially relevant for designing collectively intelligent systems. And as noted below, it may also be useful to add new moves for other types of problems.

In addition, even within a given domain, some moves may be more useful than others for specific problems. For example, we demonstrated here that the most innovative move provided by Ideator tended to differ depending on which problem statement was being examined. Although Ideator provides a wide array of moves, it does not currently tailor the moves to specific types of problems. As such, we expect one focus of our future work to involve suggesting specific moves or sequences of moves for specific problem types and situations.

6.4 Recombination of ideas

An important part of creative problem-solving is often the selection of fragments of ideas that can be combined with others. For example, our UI supports this process, in part, by letting users bookmark key results, thus reducing the potential for good ideas to get lost in the avalanche of text that LLMs can produce. We believe future interfaces should also be designed in ways that explicitly support idea synthesis and recombination, for example by rating the sub-parts of an idea (at the word or sentence level) rather than simply rating an entire idea.

6.5 Community-based extension

While we have not yet demonstrated this capability, the structure of the Ideator suggests how a specialized front end and flexible back end can help create a platform for community creation and evolution. Collective software development by an open-source community, for example, can improve or add individual "moves" or run combinations of moves within custom UIs. We hope that our contribution scaffolds future design and development of tools that augment the collective ability of many others.

7 FUTURE WORK

It appears that systems like the Supermind Ideator have the potential to provide substantial assistance to humans doing creative problem-solving. But much work remains to be done for this potential to be fully realized, and we see at least three short-term paths forward to improve Ideator:

7.1 Adding and Improving Moves

One of our short-term goals is to refine the prompts and fine-tuning of our moves to improve the quality of the ideas the moves generate. In addition, it seems very promising (and feasible) to add support

for new types of moves. Some of these moves might be relevant for creative problem-solving in any domain. For example, moves might be developed for techniques such as lateral thinking [13] or "5 W's and 1 H" [4].

It would also be possible to include moves to support other methodologies for thinking about business questions, such as Porter's 5-forces [47] or Blue Ocean Strategy [27]. And it would potentially be possible to add moves to support other specific topic domains such as mechanical design or building architecture.

We hope that by having an extensible and open-source API for the Supermind Ideator, communities of professionals, researchers, consultants, and others could contribute to a growing collective of knowledge about how to guide LLMs to help with specific types of ideation and creative problem solving.

7.2 Evaluating Ideas

Currently, the moves we have implemented only cover the divergent aspects of idea generation: generating possible ways of re-framing the problem and generating possible solutions to the problem. As the double diamond process suggests, however, evaluating and selecting among these possibilities is also necessary to be able to actually use the results.

For instance, as described above, the thumbs-up, thumbs-down, and bookmark buttons provide simple tools for evaluating ideas. With these features, users can quickly skim through their previous ideas and find things they liked, disliked, or saved.

Looking forward we feel that idea evaluation can be greatly expanded upon in a number of ways such as sharing ratings in a group and using previous ratings to recommend items in a new situation.

7.3 Adding Exploration Modes

The Supermind Ideator currently produces output in a single thread of ideas. Each round of idea generation comes after the next in a linear manner. Re-running moves allows for a nesting of output, but that is as far as we currently go. From our user evaluation we found that people do not always work in linear patterns of thought and would occasionally find themselves wanting to combine outputs from different moves or otherwise reorganize output in a manner that allowed the user to see clusters of thought and reflect on them in non-threaded and non-linear manners.

Taking into consideration the power of affinity diagrams when brainstorming, we believe a more flexible and dynamic interface could be designed to enable users to connect their input, output, and associated ideas in less restrictive ways. For example, by giving 'sticky note' like flexibility to the interface there might be increased capacity for synthesis of output.

Furthermore, if the idea generation process is thought of as traversing a graph, then an interface that treats each idea as a node could allow for one to see the spaces they have already deeply explored vs. those they have not yet tapped into and either help broaden exploration or narrow focus.

8 CONCLUSION

In this paper, we have shown how the use of generative AI technologies, like LLMs, can benefit substantially from scaffolding designed

to help users perform specific kinds of tasks. In particular, we have shown how these systems can be customized to help people do creative problem-solving in a broad range of areas and then further customized to help with creative problem-solving in the specific area of designing groups of people and computers. We believe that these approaches can also be extended to include many other creativity techniques and many other techniques for other topic areas.

In the long run, we hope that tools like this will help meet one of the most important practical challenges facing the field of collective intelligence: How can we design more collectively intelligent systems of people and computers to deal with our most important problems in business, government, science, and many other areas of society?

ACKNOWLEDGMENTS

This research was supported in part by the Toyota Research Institute and by the Collective Intelligence Mission of the MIT Quest for Intelligence. We also want to thank Hendrik Maier and Stephen Dwyer for their help with the early Ideator concept and Lilly Kammerling for her help with the Ideator API.

REFERENCES

- [1] 2005. Framework for Innovation: Design Council's evolved Double Diamond. <https://www.designcouncil.org.uk/our-work/skills-learning/tools-frameworks/framework-for-innovation-design-councils-evolved-double-diamond/>
- [2] 2018. Design Thinking Bootleg. <https://dschool.stanford.edu/resources/design-thinking-bootleg>
- [3] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 610–623. <https://doi.org/10.1145/3442188.3445922>
- [4] Tom Brown. 2009. *Change by design*. HARPER COLLINS.
- [5] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Inc., 1877–1901. <https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418fb8ac142f64a-Abstract.html>
- [6] Erik Brynjolfsson, Danielle Li, and Lindsey R Raymond. 2023. *Generative AI at work*. Technical Report. National Bureau of Economic Research.
- [7] Bendik Bygstad. 2010. Generative mechanisms for innovation in information infrastructures. *Information and Organization* 20, 3 (July 2010), 156–168. <https://doi.org/10.1016/j.infoandorg.2010.07.001>
- [8] Andres Campero, Michelle Vaccaro, Jaeyoon Song, Haoran Wen, Abdullah Almaatouq, and Thomas W. Malone. 2022. A Test for Evaluating Performance in Human-Computer Systems. <https://doi.org/10.48550/arXiv.2206.12390> arXiv:2206.12390 [cs].
- [9] Carlos Cardoso and Petra Badke-Schaub. 2011. Fixation or Inspiration: Creative Problem Solving in Design. *The Journal of Creative Behavior* 45, 2 (June 2011), 77–82. <https://doi.org/10.1002/j.2162-6057.2011.tb01086.x>
- [10] Fang Chen and Jacques Terken. 2023. Design Process. In *Automotive Interaction Design: From Theory to Practice*, Fang Chen and Jacques Terken (Eds.). Springer Nature, Singapore, 165–179. https://doi.org/10.1007/978-981-19-3448-3_10
- [11] Nigel Cross. 1982. Designerly ways of knowing. *Design Studies* 3, 4 (Oct. 1982), 221–227. [https://doi.org/10.1016/0142-694X\(82\)90040-0](https://doi.org/10.1016/0142-694X(82)90040-0)
- [12] Rikke Friis Dam and Teo Yu Siang. 2020. Design thinking: a quick overview. (June 2020). <https://apo.org.au/node/306478> Publisher: Interaction Design Foundation.
- [13] Edward De Bono. 1970. Lateral thinking. *New York* (1970).
- [14] Kees Dorst and Nigel Cross. 2001. Creativity in the design process: co-evolution of problem-solution. *Design Studies* 22, 5 (Sept. 2001), 425–437. [https://doi.org/10.1016/S0142-694X\(01\)00009-6](https://doi.org/10.1016/S0142-694X(01)00009-6)
- [15] Claudia Eckert and Martin Stacey. 2000. Sources of inspiration: a language of design. *Design Studies* 21, 5 (Sept. 2000), 523–538. [https://doi.org/10.1016/S0142-694X\(00\)00022-3](https://doi.org/10.1016/S0142-694X(00)00022-3)

- [16] Andreas Fink, Karl Koschutnig, Mathias Benedek, Gernot Reishofer, Anja Ischebeck, Elisabeth M. Weiss, and Franz Ebner. 2012. Stimulating creativity via the exposure to other people's ideas. *Human Brain Mapping* 33, 11 (2012), 2603–2610. <https://doi.org/10.1002/hbm.21387>
- [17] MIT Center for Collective Intelligence. [n. d.]. Supermind Design Primer | MIT Center for Collective Intelligence. <https://cci.mit.edu/supermind-design-primer/>
- [18] Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2018. Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 1235–1257. <https://doi.org/10.1145/3196709.3196732>
- [19] Liane Gabora. 2002. Cognitive mechanisms underlying the creative process. In *Proceedings of the 4th conference on Creativity & cognition (C&C '02)*. Association for Computing Machinery, New York, NY, USA, 126–133. <https://doi.org/10.1145/581710.581730>
- [20] K. J. Gilhooly, E. Fioratou, S. H. Anthony, and V. Wynn. 2007. Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology* 98, 4 (2007), 611–625. <https://doi.org/10.1111/j.2044-8295.2007.tb00467.x>
- [21] Steven M. Goodman, Erin Buehler, Patrick Clary, Andy Coenen, Aaron Donsbach, Tiffanie N. Horne, Michal Lahav, Robert MacDonald, Rain Breaw Michaels, Ajit Narayanan, Mahima Pushkarna, Joel Riley, Alex Santana, Lei Shi, Rachel Sweeney, Phil Weaver, Ann Yuan, and Meredith Ringel Morris. 2022. LaMPost: Design and Evaluation of an AI-assisted Email Writing Prototype for Adults with Dyslexia. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, 1–18. <https://doi.org/10.1145/3517428.3544819>
- [22] Dawn G. Gregg. 2010. Designing for collective intelligence. *Commun. ACM* 53, 4 (April 2010), 134–138. <https://doi.org/10.1145/1721654.1721691>
- [23] Matthias Griebel, Christoph Flath, and Sascha Friesike. 2020. AUGMENTED CREATIVITY: LEVERAGING ARTIFICIAL INTELLIGENCE FOR IDEA GENERATION IN THE CREATIVE SPHERE. (2020).
- [24] T. J. Howard, E. A. Dekoninck, and S. J. Culley. 2010. The use of creative stimuli at early stages of industrial product innovation. *Research in Engineering Design* 21, 4 (Oct. 2010), 263–274. <https://doi.org/10.1007/s00163-010-0091-4>
- [25] Edwin L Hutchins, James D Hollan, and Donald A Norman. 1985. Direct manipulation interfaces. *Human-computer interaction* 1, 4 (1985), 311–338.
- [26] David G. Jansson and Steven M. Smith. 1991. Design fixation. *Design Studies* 12, 1 (Jan. 1991), 3–11. [https://doi.org/10.1016/0142-694X\(91\)90003-F](https://doi.org/10.1016/0142-694X(91)90003-F)
- [27] W. Chan Kim and Mauborgne Renee. 2017. *Blue Ocean Shift: Beyond competing: Proven steps to inspire confidence and seize new growth*. Hachette Books.
- [28] Mark Klein and Gregorio Convertino. 2014. An embarrassment of riches. *Commun. ACM* 57, 11 (2014), 40–42.
- [29] Don Koberg and Jim Bagnall. 1974. The Universal Traveler, A Soft-Systems Guide to: Creativity, Problem Solving, and the Process of Reaching Goals.[Revised Edition.]. (1974).
- [30] Akhilesh Koppeneni, David Sun Kong, and Thomas W Malone. 2022. Supermind Design for Responding to Covid-19: A case study of university students generating innovative ideas for a societal problem. (2022).
- [31] Jo'Anne Langham and Neil Paulsen. 2020. Simulating the cognitive leap using brainwriting. *Design Management Journal* 15, 1 (2020), 85–101.
- [32] Robert Laubacher, Annalyn Bachmann, Kathleen Kennedy, and Thomas W Malone. 2023. Supermind Design for inventing smarter organizations: Applying a new organizational design approach in a professional services setting. (2023).
- [33] Robert Laubacher, Gianni Giacomelli, Kathleen Kennedy, David Sun Kong, Annalyn Bachmann, Katharina Kramer, Paul Schlag, and Thomas W. Malone. 2020. Using a Supermind to Design a Supermind: A Case Study of University Researchers and Corporate Executives Co-Designing an Innovative Healthcare Concept. <https://doi.org/10.2139/ssrn.3601059>
- [34] Jung-Yong Lee and Chang-Hyun Jin. 2019. How Collective Intelligence Fosters Incremental Innovation. *Journal of Open Innovation: Technology, Market, and Complexity* 5, 3 (Sept. 2019), 53. <https://doi.org/10.3390/joitmc5030053>
- [35] Terence Lee, Lauren O'Mahony, and Pia Lebeck. 2023. Prompts for Creativity. In *Creativity and Innovation: Everyday Dynamics and Practice*, Terence Lee, Lauren O'Mahony, and Pia Lebeck (Eds.). Springer Nature, Singapore, 85–115. https://doi.org/10.1007/978-981-19-8880-6_4
- [36] J. S. Linsey, I. Tseng, K. Fu, J. Cagan, K. L. Wood, and C. Schunn. 2010. A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. *Journal of Mechanical Design* 132, 4 (April 2010), 041003. <https://doi.org/10.1115/1.4001110>
- [37] Todd I. Lubart. 2001. Models of the Creative Process: Past, Present and Future. *Creativity Research Journal* 13, 3-4 (Oct. 2001), 295–308. https://doi.org/10.1207/S15326934CRJ1334_07
- [38] A. Majchrzak and A. Malhotra. 2013. Towards an information systems perspective and research agenda on crowdsourcing for innovation. *The Journal of Strategic Information Systems* 22, 4 (Dec. 2013), 257–268. <https://doi.org/10.1016/j.jsis.2013.07.004>
- [39] Thomas W Malone. 2018. *Superminds: The surprising power of people and computers thinking together*. Little, Brown Spark.
- [40] Thomas W Malone, Kevin Crowston, and George Arthur Herman. 2003. *Organizing business knowledge: The MIT process handbook*. MIT press.
- [41] Sarnoff Mednick. 1963. The associative basis of the creative process. *Psychological Review* 69, 3 (April 1963), 220. <https://doi.org/10.1037/h0048850> Publisher: US: American Psychological Association.
- [42] Cheryl Nakata and Jiyoung Hwang. 2020. Design thinking for innovation: Composition, consequence, and contingency. *Journal of Business Research* 118 (Sept. 2020), 117–128. <https://doi.org/10.1016/j.jbusres.2020.06.038>
- [43] Shakked Noy and Whitney Zhang. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381, 6654 (2023), 187–192.
- [44] Alex Osborn. 2012. *Applied imagination-principles and procedures of creative writing*. Read Books Ltd.
- [45] Paul B Paulus, Nicholas W Kohn, and Lauren E Arditti. 2011. Effects of quantity and quality instructions on brainstorming. *The Journal of Creative Behavior* 45, 1 (2011), 38–46.
- [46] Xavier Pavie and Daphne Carthy. 2015. Leveraging Uncertainty: A Practical Approach to the Integration of Responsible Innovation through Design Thinking. *Procedia - Social and Behavioral Sciences* 213 (Dec. 2015), 1040–1049. <https://doi.org/10.1016/j.sbspro.2015.11.523>
- [47] Michael E. Porter. 1998. *Competitive strategy: Techniques for analyzing industries and competitors: With a new introduction*. Free Press.
- [48] Steven R Rick, Gianni Giacomelli, Haoran Wen, Robert J Laubacher, Nancy Taubenslag, Jennifer L Heyman, Max Sina Knicker, Younes Jeddi, Hendrik Maier, Stephen Dwyer, et al. 2023. Supermind Ideator: Exploring generative AI to support creative problem-solving. *arXiv preprint arXiv:2311.01937* (2023).
- [49] Brett Robertson and David Radcliffe. 2008. The Role of Software Tools in Influencing Creative Problem Solving in Engineering Design and Education. American Society of Mechanical Engineers Digital Collection, 999–1007. <https://doi.org/10.1115/DETC2006-99343>
- [50] Elizabeth B.-N. Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *CoDesign* 4, 1 (March 2008), 5–18. <https://doi.org/10.1080/15710880701875068>
- [51] Nikhil Singh, Guillermo Bernal, Daria Savchenko, and Elena L. Glassman. 2022. Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. *ACM Transactions on Computer-Human Interaction* (Feb. 2022). <https://doi.org/10.1145/3511599> Just Accepted.
- [52] Steven M. Smith and Julie Linsey. 2011. A Three-Pronged Approach for Overcoming Design Fixation. *The Journal of Creative Behavior* 45, 2 (2011), 83–91. <https://doi.org/10.1002/j.2162-6057.2011.tb01087.x>
- [53] Hariharan Subramonyam, Christopher Lawrence Pondoc, Colleen Seifert, Maneesh Agrawala, and Roy Pea. 2023. Bridging the Gulf of Envisioning: Cognitive Design Challenges in LLM Interfaces. *arXiv preprint arXiv:2309.14459* (2023).
- [54] Douglas Summers-Stay, Clare R. Voss, and Stephanie M. Lukin. 2023. Brainstorm, then Select: a Generative Language Model Improves Its Creativity Score. <https://openreview.net/forum?id=8HwKj1wvl>
- [55] Todd M. Thrash, Emil G. Moldovan, Victoria C. Oleynick, and Laura A. Maruskin. 2014. The Psychology of Inspiration. *Social and Personality Psychology Compass* 8, 9 (2014), 495–510. <https://doi.org/10.1111/spc.12127>
- [56] Karthik Valmeekam, Alberto Olmo, Sarah Sreedharan, and Subbarao Kambhampati. 2022. Large Language Models Still Can't Plan (A Benchmark for LLMs on Planning and Reasoning about Change). *arXiv preprint arXiv:2206.10498* (2022).
- [57] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*, Vol. 30. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fb053c1c4a845aa-Abstract.html>
- [58] Joel L. Wooten and Karl T. Ulrich. 2017. Idea Generation and the Role of Feedback: Evidence from Field Experiments with Innovation Tournaments. *Production and Operations Management* 26, 1 (2017), 80–99. <https://doi.org/10.1111/poms.12613>
- [59] Robert J. Youmans and Thomaz Arciszewski. 2014. Design fixation: Classifications and modern methods of prevention. *AI EDAM* 28, 2 (May 2014), 129–137. <https://doi.org/10.1017/S0890060414000043> Publisher: Cambridge University Press.
- [60] Jinhyo Joseph Yun, Euiseob Jeong, Sangwoo Kim, Heungju Ahn, Kyunghun Kim, Sung Deuk Hahn, and Kyungbae Park. 2021. Collective Intelligence: The Creative Way from Knowledge to Open Innovation. *Science, Technology and Society* 26, 2 (July 2021), 201–222. <https://doi.org/10.1177/09717218211005604> Publisher: SAGE Publications India.
- [61] Qihao Zhu and Jianxi Luo. 2023. Generative Transformers for Design Concept Generation. *Journal of Computing and Information Science in Engineering* 23, 4 (Jan. 2023). <https://doi.org/10.1115/1.4056220>