Augmenting People's Creative Idea Generation Using an Artificial Intelligent Sketching Collaborator

Joseph Maloba Makokha

Abstract**—**Idea generation is an important part of the design process, and many strategies to support this stage have been developed. As artificial intelligence (AI) gains adoption in many domains, we need to understand its role, if any, in the design process. This paper introduces the concept of a "Disruptive Interjector", an AI system that frequently interjects with suggestions based on observing what a user does. The concept emanates from a study that was conducted with pairs of humans on one hand, and human-AI pairs on the other collaborating on idea generation by sketching. Results from a study show that participants who collaborated with, and took cues from the AI sketch suggestions generated more ideas; and also had more ideas ranked by experts as "creative" compared to two humans working together on the same tasks. It is notable that while researchers from diverse fields of engineering, psychology, art and others have explored conditions and environments that enhance people's creativity - and have provided insights on creativity in general - there still exists a gap on the role that AI can play on creativity. We attempt to narrow this gap.

Keywords**—**Artificial intelligence, design collaboration, creativity, human-machine collaboration, machine learning.

I. Introduction

THIS research seeks to understand ways in which AI can contribute to the generative part of design activity, and was conducted by involving pairs of humans on one hand, and human-AI pairs on the other, to collaborate on idea generation. In many settings, design tasks involving idea generation are often performed by teams of two or more people working synchronously, sometimes resulting in creative ideas. We also find that researchers from diverse fields such as engineering [1], psychology [2], art [3] and others have explored conditions and environments that enhance people's creativity - defined for example as their ability to come up with new, surprising, and valuable ideas [3], or finding solutions that are far from obvious through a shift in problem formulation [4]. While psychologists disagree [2] on the definition of creativity, attempts to define it straddle three aspects relating to the process Gestalt position based on insight and productive thinking that arises when one grasps the essential features and their relationship to the solution, as proposed by Wertheimer [5]; the person as emphasized by psychologist J.P. Guilford [6]; and characteristics of a product such as producing "effective surprise" combined with a "shock of recognition" as viewed by Bruner [7]. For this discussion, we will consider creativity from a product lens, where we will evaluate the outcome of design activity presented through sketching. A creative result or product is therefore one that has a level of "uncommonness" compared to others in the group, in addition to being adaptive to reality as characterized by Barron [8]; or as put differently by Newell et al. [9], one that demonstrates novelty and appropriateness or value.

Given these definitions of creativity, we see the reason why technologies like AI offer viable opportunities to augment human abilities on thinking tasks like ideation and others, something long envisioned by scholars, practitioners and researchers such as Licklider [10] who suggested that future human-machine partnerships would perform intellectual operations more effectively than a human alone; and Simon [11], who anticipated machines that would be capable of a range of thinking tasks similar to those that humans perform. In today's work environment with an increasing presence of AI agents, humans are frequently collaborating with machines in Human-AI (H-AI) teams in diverse fields ranging from medicine [12], [13] where they aid in detection of fractures and tumors [14]; to data science where they automate tasks; and in robotics and self-driving cars [15] among others.

When it comes to collaboration on designing, the process goes through a phase where ideas are generated, considered from multiple points of views, and a final candidate is chosen. This process moves from a divergent (idea generation) to a convergent one where the best design is selected. It is generally desired to quickly generate multiple ideas of high quality [1], and then efficiently select the best. In line with this possibility of good ideas, the generative phase has proved valuable in many other fields beyond design. In engineering economics for example, innovative alternatives often have far greater benefits and payoffs than more accurately analyzing existing alternatives [16]. We consider a common scenario where choosing between two alternatives whose costs differ by 10% may require accurate estimates and detailed analysis - all to achieve a maximum payoff of 10% - while a new alternative that is 50% better can be justified and implemented with less detailed analysis, because it is clearly superior, demonstrating the advantages of idea generation.

II. Related Work

Technology-Mediated Collaboration

Applying technology in ideation has been explored by researchers in studies where design activity is mediated with technology. In past studies Bamber [17]; Karan et al. [18]; Murthy and Kerr [19] have compared outcomes in problem solving tasks from groups using computer-mediated

Joseph Maloba Makokha is with Stanford University, USA (e-mail: makokha@stanford.edu).

communication (CMC) to those working unassisted/face-to-face (FTF), and found that CMC brainstorming generated more ideas (divergence) than FTF while unassisted teams weeded out irrelevant ideas and recommendations (hence converged) better than CMS teams [20]. While these studies demonstrate the benefits of using computing technologies in idea generation, these technologies are applied differently since they are used to manage flow of information, while we are applying technology to influence the kind of ideas that are conceived by a person. Others have set out a framework and developed a textual design assistant called the Problem Formulator [21] that takes in specifications and requirements, then provides a set of designs, as an example of computational tools that help designers formulate the problem (rather than solving the problem) during early concept development. This tool therefore differs in application from our context even though it is suitable for engineering design - due to its support for distinct alternatives incorporating multiple protosolutions.

In other studies examining characteristics of divergent and convergent information types in product development teams, Schar [20] identifies representational gaps - arising from individual team members' cognitive preferences. He proposes “pivotal thinking” as a bridging mechanism, which suggests that technologies such as AI might help bridge such representational gaps in order to provide coherence between team members, especially in situations where multiple members collaborate with technology. Another factor considered in past studies is the effect on the mean judgment (decision) of team members working FTF compared with a computer system (Group Decision Support System), where Karan et al. [18] found a significant cautious shift (change from judgment on the same task when performed individually) in the FTF teams, and no significant shift in the computer-mediated groups. This suggests benefits of teamwork, such as enabling people to come up with ideas that they consider better than the ones created when they work alone, among other benefits.

Traditional (Rule-Based) AI and Big Data Approaches

Both big data as well as rule-based systems have demonstrated promising results. One interesting approach can be seen in efforts to build medical AI systems [22] that create human-observable structures, thereby providing heuristics from which humans doctors can learn. This is achieved by applying a method that draws from shape grammars and graph grammars, which are rule-based techniques in design and architecture. To implement this system, they added a preprocessing step to turn each individual angiogram into many spatially derived features that enable application of machine learning on the data. Notably, they used relatively small datasets to discover indicative rules that help in detecting anomalies in vascular conditions with high accuracy.

Sketching, Visualizing, Reinterpreting

People demonstrate different preferences in their thinking styles, though sketching is common among designers in expressing ideas from early thoughts to more elaborate features and details. It turns out that sketching offers distinct advantages to design activity. For instance, Tversky et al. [23] suggest that sketching helps to relieve short-term memory, establish consistency, and augment information processing just like other external representations that we use in design. They add that since these diagrams tend to be “sketchy”, in other words, vague, committing only to minimal global configurations and sizes, they are full of ambiguities - especially so for early sketches - and rather than creating a source of uncertainty and confusion, such ambiguity in design sketches can be a source of creativity because it allows perceiving and reinterpreting figures and groupings of figures. Further, they observe that while both experienced as well as novice designers are able to make new inferences from their own sketches, experienced designers are more adept at making functional inferences (such as seeing the flow of pedestrians in a sketch of a building’s layout, or chess pieces’ motion on a real or imaginary chessboard), compared with novices whose inferences are primarily perceptual (like seeing new spatial relations among structures).

In one of their studies Tversky et al. [23] tested novice designers (undergraduates) to see if searching for new perceptual relations could be deliberately used to enable new interpretations by a larger population. Within the participating students, two broad categories emerged. Given 4 minutes of time, the participants who adopted a strategy (either spontaneously or by suggestion) of attending to the parts of the sketch, either by focusing on different parts or mentally rearranging the parts of the sketch to find new interpretations generated more interpretations (45 average for the different parts group, 50 for rearrange parts group), than those who did not adopt the strategy (27 average). Such results suggest that providing designers with cues that encourage reinterpretations of a task might lead to more ideas, and higher chances of good ideas [1].

Measuring Ideation Effectiveness

When it comes to measuring ideation effectiveness, researchers like Shah et al. [24] have proposed methods that can be applied in classifying design sketches produced as solutions to specific prompts (solution-focused ideation) involving a measure of the output and its function. For this study, tasks conveying an output (e.g. ‘a way to get a toddler into bed’) as well as a set of functional requirements (e.g. have a baby inside a bed), created by participants in a controlled study are evaluated. Expert ratings are another viable approach, though Shah et. al.’s systematic approach with its methodical, consistent procedures has the benefits of producing more comparable and replicable results [25].

Given the foregoing insights on collaborative design activities, some involving humans and machines, we seek to answer the following questions:

- How can AI contribute to the generative part of the design process? Further:
- Given that machines have no sense of understanding or contextual details (like humans do), how do they perform as partners for divergent thinking design tasks? (We consider quantity and quality of ideas)

- How are they perceived as partners for divergent thinking design tasks? (We ask participants how it was like to have a partner that is an AI system rather than a human design partner)

The Disruptive Interjector: A Distinct Class of AI System

We explore a class of AI systems that track human designers' actions (e.g., sketches) and interject occasionally with comments that might lead the designers down promising new paths. This is just one of the many ways that AI systems might augment human abilities in the design process. We therefore coined the term "Disruptive Interjector" (DI), to differentiate this class of computational artifacts from others that function at roughly the same level of abstraction, for example the "recommender system" or "cognitive tutor". In other words, this is an interactive "hint giver or AI Collaborator" that is distinct from "chat bots/voice assistants", hence the name DI that characterizes the unsolicited interaction initiated by the system during collaboration between a designer and this AI system.

There are several related but different concepts that work at the same conceptual level as the DI system we are exploring. These comprise Cognitive Tutors; Recommender Systems; Priming and Anchoring; Planning Systems; and Design Systems. Here is a brief distinction to differentiate these systems, also summarized in Tables I and II:

- A Cognitive Tutor needs to command a higher level of expertise relative to the learner, while this generative AI tool acts as a design partner that may have about the same or lower level of knowledge than a human.
- The Recommender System needs access to previous selections to determine preferences, then make new suggestions. While these two systems rely on a user's previous selections/responses in a specific context, the generative AI tool has limited contextual understanding of the task, and attempts to predict sketches as they develop, then suggest new images that may or may not be related to the observed sketch, and which sometimes lead the designer to new ideas.
- Priming/Anchoring; Planning Systems; and Design Systems rely on an independent source of input rather than a designer/user's selections as opposed to the DI, Cognitive Tutor or Recommender Systems, which make decisions based on a designer/user's actions.

Notably, the generative AI system aims to diverge, as opposed to the others which converge towards a solution or answer. Such divergence is desirable during idea generation, given that it may help resolve "idea-fixation" [26], [27] that tends to inhibit a consideration of other possibilities.

III. Methodology

This study involved 29 participants in two categories where 17 (59%) participants were each paired with a computer system to form Human-AI (H-AI) pairs, while the remaining 12 (41%) collaborated in 6 teams of human-human (H-H) pairs. Within these participants, 22 (76%) were men and 7 (24%) were female. All were novice designers enrolled in cross-disciplinary programs that entailed a mix of domains including psychology and engineering as advanced undergraduates or beginning graduate students.

TABLE I
Differences between AI Generative Partner (DI) and Other Systems that Observe/Track User Input, then Offer Suggestions

Tool	Goal	Method	Direction
DI	Lead a novice designer to imagine new ideas	Tracks a user's actions and interjects with suggestions	Divergence
Cognitive Tutor	Guide the learner based on their performance or level of understanding	Monitors actions, then modifies subsequent content to fit learning goals (for routine procedures)	Diverge/ Converge
Recommender System	Offer the user one or more options from a larger collection, based on their previous choices	Uses desired preferences to suggest likely candidates that meet the requirements (for simple choice tasks e.g. movie to watch)	Convergence

TABLE II
Differences between AI Generative Partner (DI) and other systems that Offer Suggestions without Needing a User's Input

Tool	Goal	Method	Direction
Priming/ Anchoring	Influence outcome to a desired one by offering a strategic starting point	Use a given form of beginning information to influence a user's decisions	Convergence
Planning System	Use available information to narrow down to a solution that offers the best outcomes from a complex set of options	Applies metrics to characteristics of the system to find the combination that results in an outcome with the greatest benefits	Convergence
Design System	Optimally identify one or more solutions from a large set of options, based on a preferred criterion	Minimizes the cost between starting and final points	Convergence

Participants received general instructions on the tasks, as well as orientation on how to use the AI web application on a tablet computer. They then collaborated on two design tasks, one after another for 7 minutes, with the difference that the AI sketching assistant was disabled for the H-H pairs. Each task involved brainstorming of ideas followed by a selection of one final choice from the generated ideas. They designed something for another user in the first task, while the second one involved designing something for themselves. At the end of the sketching tasks, a questionnaire was administered, and an exit interview conducted to learn about their thought process around strategies they applied when collaborating with either another human, or an AI agent instead of a human. Other information collected included: Team demographics (age, level of design training, cultural background, and their affinity to teamwork); Start and end time of each round; Number of sketches, and their timestamps; and whether they collaborated with a human or AI.

Developing Prompts to Use in the Study

The prompts used for the design activities were crowdsourced from a group comprising a mix of novice and experienced engineering designers who were asked to list some meaningful but challenging tasks they had recently encountered in the course of daily lives. They subsequently generated over two dozen tasks (hereafter referred to as “needs”) of which 10 are listed in Table III. These 10 needs were selected because they were deemed accessible to the prospective study participants (novice engineering design students) derived from university students.

TABLE III
TASKS AND ANTICIPATED LENGTH OF TIME REQUIRED

Design Tasks	Task Length
Water bottle for elderly person	short
Way for a toddler to get onto their bed	short
Way to transport your laptop	long
Way to secure a bike on campus	long
Keychain that is easy/obvious to locate	short
Cat-activated water dispenser for cat to drink from	short
Toe nail clipper for someone who cannot reach their toes	short
Car handle for elder with grasp problem	short
Collapsible bike helmet	short
Safe raincoat for cyclist to wear while riding bikes	short

Of the ten prompts above, two were selected as suitable because of their level of abstraction, relevance to most people, and flexibility for modification or reframing. These were to design a way to:

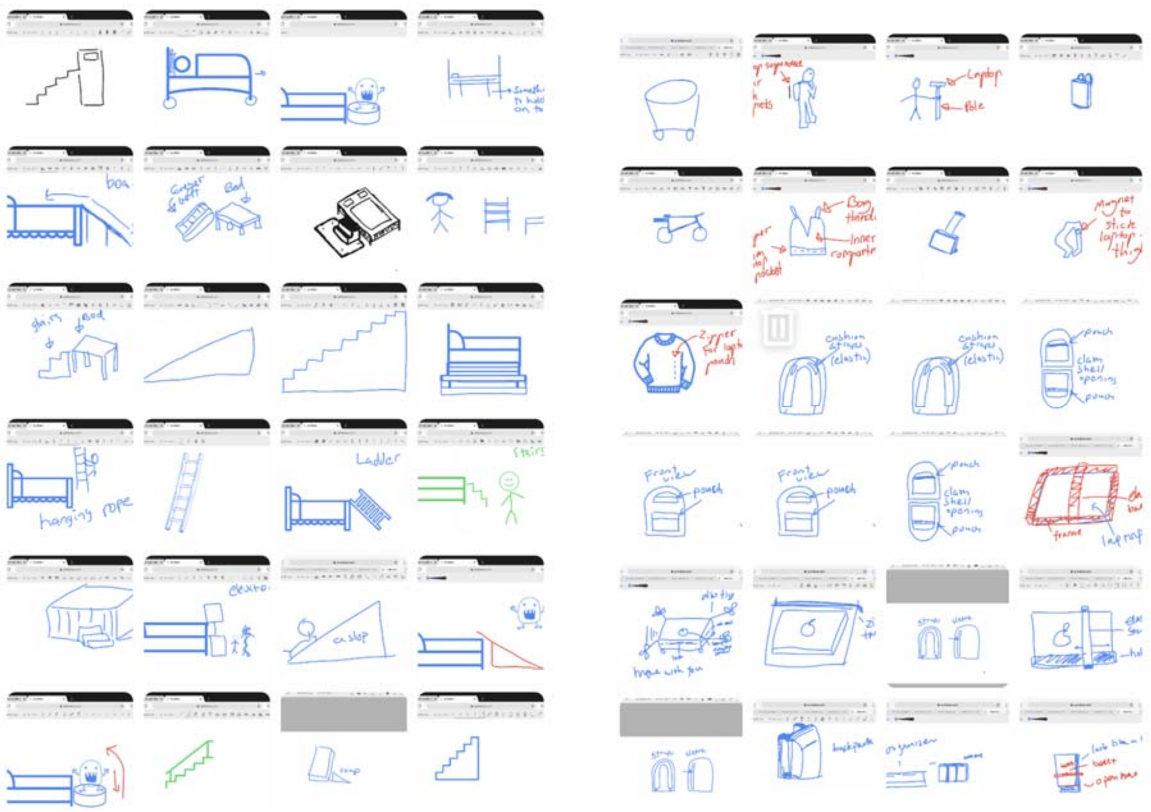
- get a toddler into bed
- transport your computer on campus

Data Collection Process

Once the participants received an introduction to the study and an orientation on using the AI sketching app, including how to capture each idea as a screenshot, they were given the first prompt and asked to proceed on the timed task. They each had a stylus (for H-H pairs, while H-AI pairs needed a single stylus) with which to draw on a tablet running a web application of the AI sketching app. They were alerted halfway through the time, as well as a minute to the end of their task. The second task followed immediately after this, with a reading of the prompt, as well as a repeat of instructions on how to capture each idea.

At the end of the two sketching activities, every participant individually completed a questionnaire in a private space; and was then briefly interviewed about the strategies they used in coming up with ideas, as well as those used in selecting their best idea. They were free to talk about their experience collaborating with another human, or with an AI agent - which yielded several insights. Consequently, we collected the following data:

- Sketches (8 to 15 per team)
- Survey (1 per participant)
- Audio of debriefing interview (1 per participant, ranging from 1 to 5 minutes)



(a) Ways of getting a toddler into bed

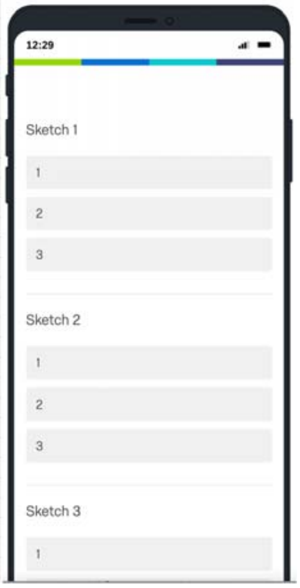
(b) Ways of transporting your computer on campus

Fig. 1 Sample sketches for the two tasks

Each of these anonymized data were captured with a date and time stamp due to the nature of digital capturing devices used in the processes.

BED_ALL

SKETCH NO.	SCORE	DATE STAMP	TIME STAMP
1		'Jul 17	2 07 49 PM.png'
2		'Jul 25	3 49 07 PM.png'
3		'Jul 17	2 06 49 PM.png'
4		'Jul 24	12 24 24 PM.png'
5		'Jul 24	12 23 52 PM.png'
6		'Jul 24	12 27 08 PM.png'
7		'Jul 26	1 17 51 PM.png'
8		'Feb 20	7 38 40 AM.png'
9		'Jul 19	10 54 23 AM.png'
10		'Jul 18	10 15 30 AM.png'
11		'Jul 26	1 14 44 PM.png'
12		'Mar 02	7 39 33 AM.png'
13		'Jul 19	10 11 37 AM.png'
14		'Jul 19	9 10 11 AM.png'
15		'Jul 19	10 17 33 AM.png'
16		'Jul 30	12 24 36 PM.png'
17		'Jul 30	12 25 36 PM.png'
18		'Jul 30	12 21 44 PM.png'
19		'Jul 30	12 20 22 PM.png'
20		'Jul 09	2 23 16 PM.png'
21		'Jul 19	9 15 07 AM.png'
22		'Feb 20	3 18 21 PM.png'
23		'Jul 19	10 10 37 AM.png'
24		'Jul 24	3 19 58 PM.png'
25		'Jul 17	3 45 27 PM.png'
26		'Jul 24	3 28 58 PM.png'
27		'Jul 24	3 16 57 PM.png'
28		'Jul 26	1 19 56 PM.png'
29		'Jul 16	9 33 AM.png'
30		'Jul 25	3 38 09 PM.png'
31		'Jul 19	10 14 09 AM.png'
32		'Jul 19	10 52 32 AM.png'
33		'Jul 09	2 24 17 PM.png'
34		'Jul 20	3 33 36 PM.png'
35		'Jul 20	3 34 08 PM.png'
36		'Jul 19	10 15 07 AM.png'



(a)

	SKETCH NO.	SCORE	Condition	DATE STAMP	TIME STAMP
TEAM1	22	1	H-H	Feb 20	3 18 21 PM.png
	44	2	H-H	Feb 20	3 19 00 PM.png
	45	2	H-H	Feb 20	3 20 58 PM.png
	8	1	H-H	Feb 20	7 38 40 AM.png
TEAM2	20	1	HAI	Jul 09	2 23 16 PM.png
	33	1	HAI	Jul 09	2 24 17 PM.png
	49	2	HAI	Jul 09	2 24 31 PM.png
	64	2	HAI	Jul 09	2 25 04 PM.png
	65	2	HAI	Jul 09	2 25 41 PM.png
	50	3	HAI	Jul 09	2 26 15 PM.png
TEAM3	39	3	HAI	Jul 17	2 06 03 PM.png
	3	1	HAI	Jul 17	2 06 49 PM.png
	1	1	HAI	Jul 17	2 07 49 PM.png
	37	2	HAI	Jul 17	2 09 25 PM.png
TEAM4	25	1	HAI	Jul 17	3 45 27 PM.png
	59	2	HAI	Jul 17	3 46 48 PM.png
	60	2	HAI	Jul 17	3 47 45 PM.png
TEAM5	66	1	HAI	Jul 18	10 13 20 AM.png
	10	1	HAI	Jul 18	10 15 30 AM.png
	51	2	HAI	Jul 18	10 16 05 AM.png
	54	2	HAI	Jul 18	10 16 44 AM.png
	52	3	HAI	Jul 18	10 18 27 AM.png
	67	3	HAI	Jul 18	10 22 52 AM.png
	53	3	HAI	Jul 18	10 23 28 AM.png

(b)

Fig. 2 Ranking of the sketches

IV. Exemplar Dataset and Analysis

The entire collection of sketches was first checked for any errors and anomalies, a step that revealed bad data from teams that failed to follow the instructions, thereby sketching multiple ideas in a single screen. Other anomalies were multiple screenshots of the same sketch, and marks from trying out features of the AI sketching app. Such sketches were eliminated from ranking. Fig. 1 presents a random sample of sketches for the two tasks, taken from a collection of approximately 300 sketches.

The sketches were evaluated by an expert and classified as follows:

- Not creative - score of 1
- Somewhat creative - score of 2
- Creative - score of 3

In order to ensure that the ranking could be replicated, we validated it using two separate processes.

Validating the Ranking Scale

The first validation of ranking involved three independent expert evaluators who ranked the sketches based on whether the sketched idea could fulfill the task, as well as its novelty. Their results matched those from the earlier ranking by the first expert. In order to increase reliability of the process, the presentation of the sketches was deliberately modified - the order of the sketches was scrambled so that evaluators saw sketches from different teams appearing randomly. That way, the evaluation was not team by team, but rather sketch by sketch, in order to minimize any bias likely to arise from evaluating one team's sketches at a time. Fig. 3 shows a sample of the unranked list given to evaluators (and a picture showing how it was displayed on a mobile device screen) in Fig. 2 (a); and a completed one arranged by team with the score, condition, date/time stamps all visible as seen in Fig. 2 (b).

From the first validation using 67 images, 100% of all images previously ranked as 1 (not creative), or 2 (somewhat creative), received a score of 1 or 2 from the expert validation ranking.

The second validation method was by using an AI classifier (Google's Teachable Machine), which was trained from a smaller sample of 36 images that received the same rank across the three evaluators, then tested with 10 unseen images from the same distribution as the training sample. It produced similar results to those of the experts when shown these test sketches, and at a 100% accuracy. In Fig. 3 are two examples of the machine's output.

V. Analysis

First Task: A Way to Get a Toddler into Bed

The results from 15 teams (4 H-H and 11 H-AI) out of the 23 were analyzed - the other 2 of the H-H and 6 H-AI require additional processing since they either created complex sketches or failed to capture images, which would require additional steps such as evaluating ideas using the video captured during their sessions. A complex sketch example is one where multiple (and different) ideas are presented in one screenshot as opposed to capturing each idea separately. Fig. 4 shows scores for all the sketches from the first task (getting a toddler to bed), grouped by team.

A graph of the scores reveals that H-H teams generated ideas that received either a "not creative" or "somewhat creative" score, while 7 out of the 11 H-AI teams (64%) came up with at least one idea that was ranked as "creative". Team 5 (H-AI) was an anomaly, generating more "creative" ideas in addition to the highest number of ideas of any team. These same results are graphed individually for the 15 teams for easier visibility in Fig. 5.

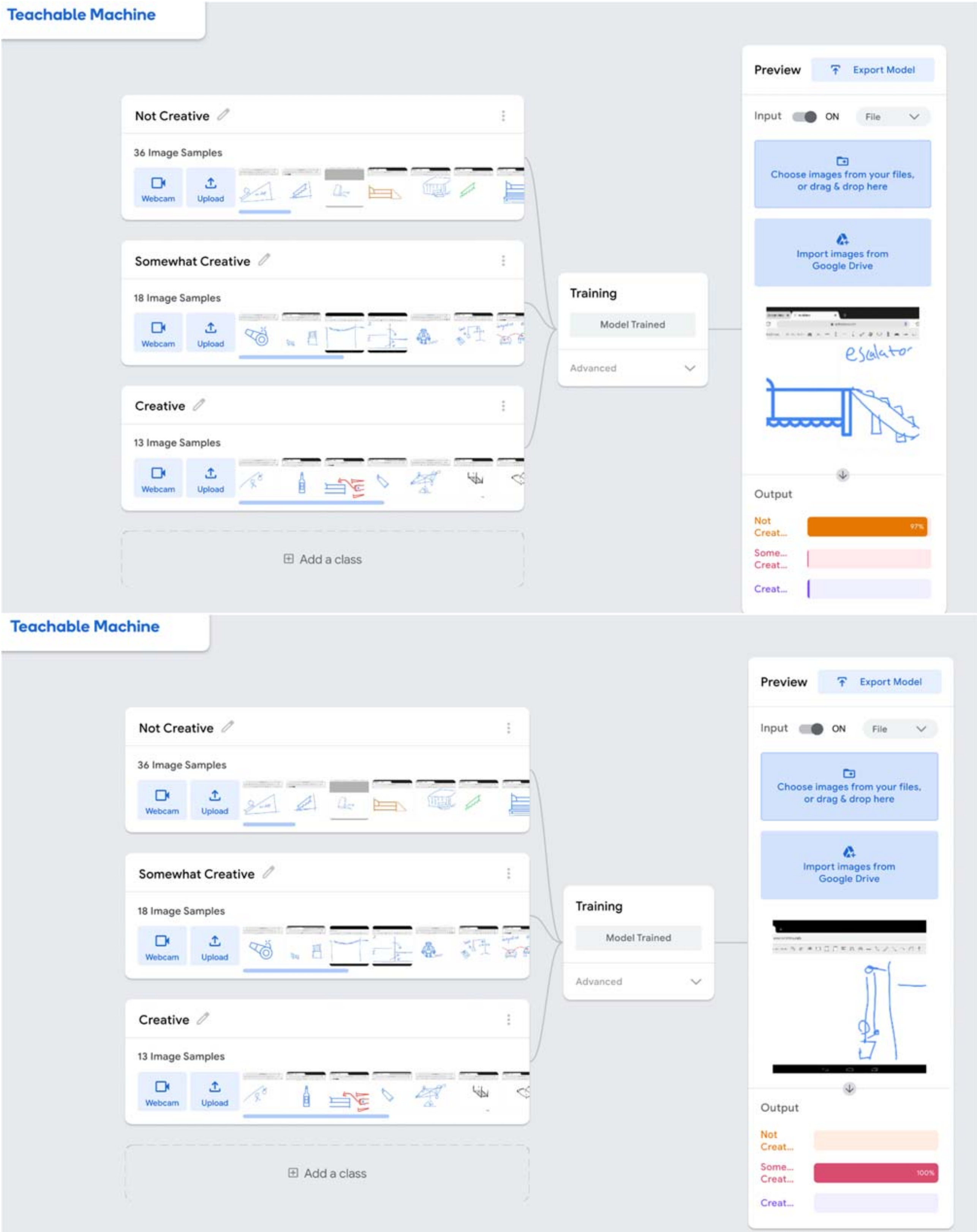


Fig. 3 Sample output from an Image Classifier

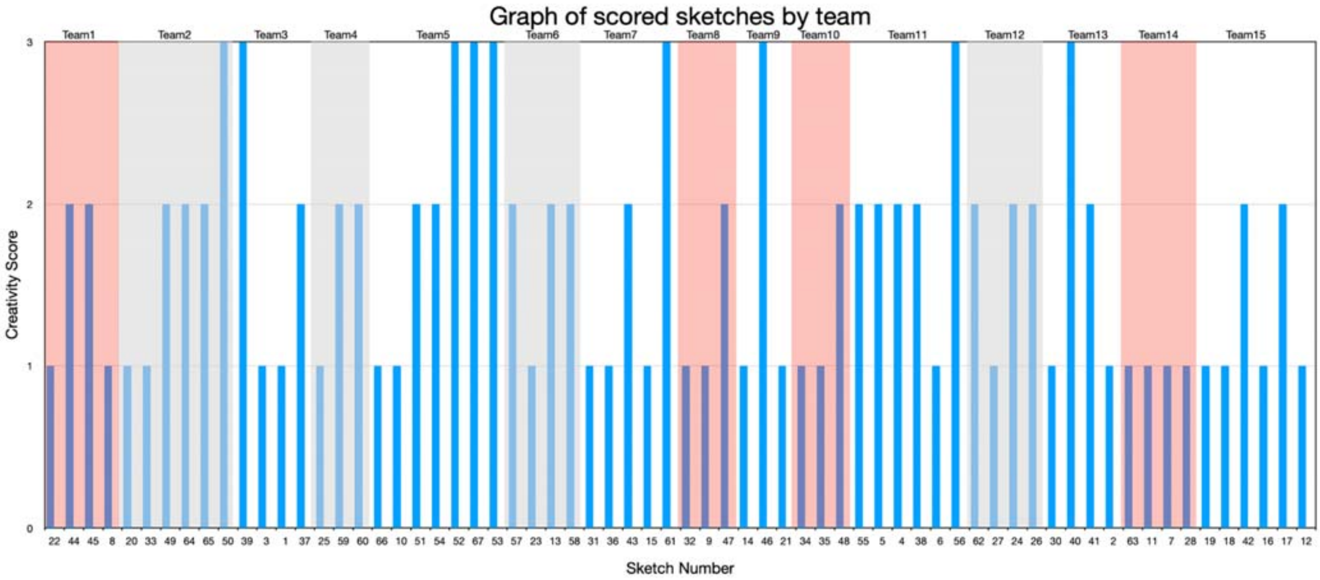


Fig. 4 Graph of scores for the 15 teams

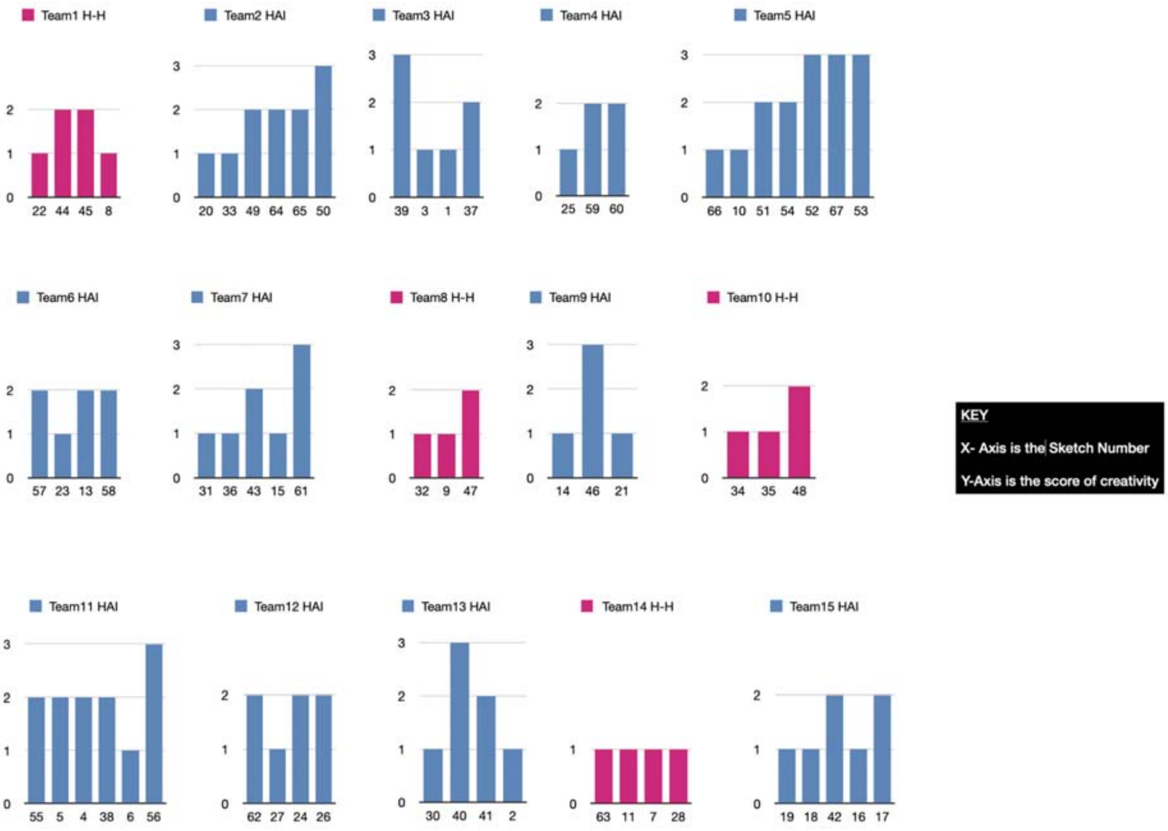


Fig. 5 Graphs of scores by team

Next in Fig. 6 is a graph combining the total number of sketches per team and a count of those ranked as not creative (NC), somewhat creative (SC) or creative (C). Of all the teams with 5 or more ideas, only one (Team 15) failed to get a "creative" rank, while for teams that generated less than 4 ideas, only one (Team 9) earned a "creative" rank on at least one idea.

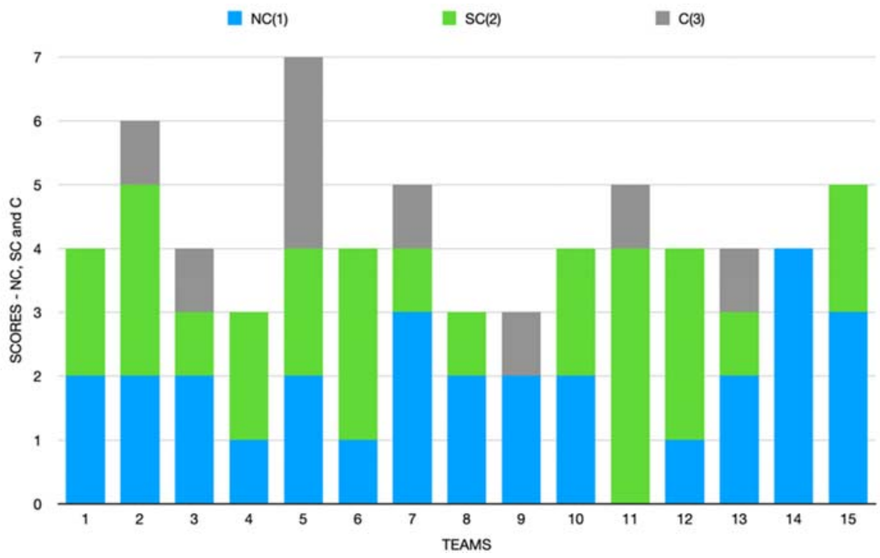


Fig. 6 Number and creativity score of sketches for each of the 15 teams

A simplified version of average number of sketches across the 15 teams, as well as for each of the ranks - not creative (NC), somewhat creative (SC) or creative (C), is shown in Fig. 7. The average number of sketches was 4.33, while that of sketches with a ranking of "not creative" was 2.0. Similarly, "somewhat creative" among the sketches averaged 1.73 (~2) while "creative" sketches averaged 0.6 (< 1). It should be noted that only 7 out of the 15 (47%) teams came up with an idea that was ranked "creative"; with none of the H-H teams achieving this rank.

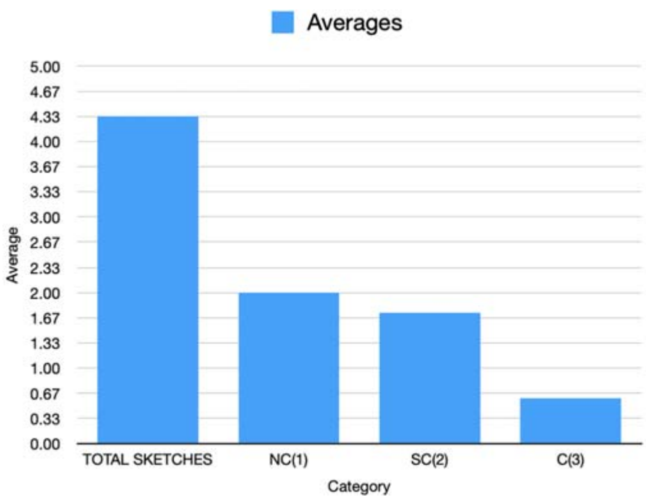


Fig. 7 Averages for number of sketches and scores for combined teams

Second Task: A Way to Transport Your Computer on Campus)

The results from this task are excluded from analysis, as they could not be validated using the two approaches employed in the first task. They require additional processing and will be included in follow up work.

VI. Approach Limitations

The first limitation we have identified is the lack of a direct method to capture how the AI generated sketch suggestions affect a participant's thought process as they conceive ideas to solve the task. We rely on the outcome of the interaction rather than an in-the-moment effect. Further, we wait until the end of the activity to debrief the participants about their collaboration, relying on their memory to recall what it was like rather than getting a sense of their interaction while it happened. Finally, without full validation of both tasks used in this study, the results from our analysis are still in the preliminary stage.

VII. Discussion and Future Directions

Evaluating Number of Sketches and Their Ranking

Teams of humans collaborating with AI had a higher number of sketches than H-H teams. The average was 4.6 for H-AI compared with 3.5 for H-H teams. This represents an average difference of 1 sketch between H-H and H-AI teams. Given that both quantity and quality matter for creative idea generation, this points to advantages derived from collaborating with an AI agent relative to another human. None of the H-H teams (which had fewer sketches on average than H-AI teams) produced an idea ranked as "creative", compared with half of the H-AI teams. This corresponds with the foregoing point on the relationship between number of ideas and the potential for "good" ideas.

The significance of the result from this study is that an artificial intelligent (AI) sketching collaborator can support ideation on divergent thinking design tasks as well as a human collaborator, and even result in more ideas generated, with a higher chance of a "creative" idea compared to teams comprising only humans. This is important since AI technologies are becoming pervasive, hence the desire to deploy these in different domains including engineering design. Based on the results from this study, we now have empirical evidence showing that AI systems that observe a person's actions, then make suggestions in the style of disruptive interjections (hence the DI phrase), can improve creativity on divergent thinking idea generation. Finally, the cross-validation from experts and the Teachable Machine (Machine Learning) demonstrate one possible way of evaluating sketches from large samples (such as in a class/course with dozens of students who sketch multiple ideas in a given session). All that would be required is for an expert to rank a few sketches that are used to train the machine learning (ML) model.

Regarding Teams of 3 or More

As for future work, the ranking of the results from the second study needs to be validated, and a larger sample size used to generate additional data to improve the machine learning validation. In addition to this, we note that engineering design teams often comprise 3 or more designers working together, hence future studies may explore such team structures and ways in which they can benefit from AI.

VIII. Conclusions

This paper presented a way in which AI might contribute to the idea generation part of design activity. We introduced the concept of a "Disruptive Interjector", an AI system that observes a person's actions, then makes suggestions based on what the user does. We demonstrated through a study conducted with pairs of humans, and human-AI pairs collaborating in idea generation, that on the basis of the number of ideas generated and those ranked as "creative", the AI collaborator worked as well or even better as a collaborator compared to two humans working together on the same tasks.

References

[1] Arnold, J. E. (1956). Creativity in engineering. SAE Transactions, 17-23.
[2] Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. Journal of personality and social psychology, 45(2), 357.
[3] Boden, M. A. (2010). Creativity and art: Three roads to surprise. Oxford University Press.
[4] Langley, P. (2018). Planning systems and human problem solving. Advances in Cognitive Systems, 7, 13-22.
[5] Wertheimer, M., & In Sarris, V. (2020). Max Wertheimer productive thinking. Cham: Springer.
[6] Guilford, J. P. (1950). Creativity. The American Psychologist, 5(9), 444–454. https://doi.org/10.1037/h0063487
[7] Bruner, J. The conditions of creativity. In Gruber, H. E., In Terrell, G., In Wertheimer, M., & University of Colorado (Boulder campus). (1962). Contemporary approaches to creative thinking: A symposium held at the University of Colorado.
[8] Barron, F. (1955). The disposition toward originality. The Journal of Abnormal and Social Psychology, 51(3), 478.
[9] Newell, A., Shaw, J. C., & Simon, H. A. (1962). The processes of creative thinking. In Contemporary Approaches to Creative Thinking, 1958, University of Colorado, CO, US; This paper was presented at the aforementioned symposium. Atherton Press.
[10] Licklider, J. C. (1960). Man-computer symbiosis. IRE transactions on human factors in electronics, (1), 4-11.
[11] Herbert A. Simon, & Allen Newell. (1958). Heuristic Problem Solving: The Next Advance in Operations Research. Operations Research, 6(1), 1.
[12] Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., ... & Potter, H. (2018). Deep neural network improves fracture detection by clinicians. Proceedings of the National Academy of Sciences, 115(45), 11591-11596.
[13] Curioni-Fontecedro, A. (2017). A new era of oncology through artificial intelligence. ESMO open, 2(2).
[14] Drozdal, J., Weisz, J., Wang, D., Dass, G., Yao, B., Zhao, C., ... & Su, H. (2020, March). Trust in AutoML: exploring information needs for establishing trust in automated machine learning systems. In Proceedings of the 25th International Conference on Intelligent User Interfaces (pp. 297-307).
[15] Yaqoob, I., Khan, L. U., Kazmi, S. A., Imran, M., Guizani, N., & Hong, C. S. (2019). Autonomous driving cars in smart cities: Recent advances, requirements, and challenges. IEEE Network, 34(1), 174-181.
[16] Eschenbach, T. (2011) "Engineering Economy: Applying Theory to Practice" Oxford University Press. Sec. 1.3
[17] Bamber, E. M., Watson, R. T., & Hill, M. C. (1996). The effects of group support system technology on audit group decision making. Auditing, 15, 122-134.
[18] Karan, V., Kerr, D. S., Murthy, U. S., & Vinze, A. S. (1996). Information technology support for collaborative decision making in auditing: An experimental investigation. Decision Support Systems, 16(3), 181-194.
[19] Kerr S. David, Murthy S. Uday Divergent and Convergent Idea Generation in Teams: A Comparison of Computer-Mediated and Face-to-Face Communication. Group Decision and Negotiation, July 2004 V13 No.4, pp 381-399
[20] Schar, M. F. (2011). Pivot thinking and the differential sharing of information within new product development teams. Stanford University
[21] MacLellan, C. J., Langley, P., Shah, J., & Dinar, M. (2013). A computational aid for problem formulation in early conceptual design. Journal of computing and information science in engineering, 13(3).
[22] Whiting, M., Mettenburg, J., Novelli, E., LeDuc, P., & Cagan, J. (2022). Inducing Vascular Grammars for Anomaly Classification in Brain Angiograms. Journal of Engineering and Science in Medical Diagnostics and Therapy.
[23] Tversky, B., Suwa, M., Agrawala, M., Heiser, J., Stolte, C., Hanrahan, P., ... & Haymaker, J. (2003). Sketches for design and design of sketches. In Human behaviour in design (pp. 79-86). Springer, Berlin, Heidelberg.
[24] Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. Design studies, 24(2), 111-134.
[25] Hay, L., Duffy, A. H., Grealy, M., Tahsiri, M., McTeague, C., & Vuletic, T. (2020). A novel systematic approach for analysing exploratory design ideation. Journal of Engineering Design, 31(3), 127-149.
[26] Leifer, L., & Meinel, C. (2019). Looking further: design thinking beyond solution-fixation. In Design Thinking Research (pp. 1-12). Springer, Cham.
[27] Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical studies of designer thinking: past, present, and future. Journal of Mechanical Design, 137(2), 021101.