

Towards Human-Robot Co-Creative Collaboration Through Interactive Task Dialogue

Hong Lu¹[0009-0004-2194-0277], Jingwen Feng¹, Emma Bethel¹, Vasanth Sarathy¹[0000-0003-4129-6429], Elaine Short¹[0000-0002-4789-9979], and Matthias Scheutz¹[0000-0002-0064-2789]

Tufts University, Medford, MA 02155, USA. <https://www.tufts.edu/>

Abstract. There is currently relatively little work on architectures and evaluations of robots that support creative designs of human interactants in manipulation tasks through dialogue. We build a dialogue system by integrating a large language model into a robot cognitive architecture and investigate whether engaging in dialogues with the robot collaborator on a creative task has positive effects on task satisfaction and experience of the human, evaluating the effect of dialogue on perceived robot utility, intelligence, and supportiveness as a creative partner. We test our hypotheses with a cake decoration task during which participants collaborate with a robot arm equipped with a mixed-initiative co-creative dialogue system to place decorations on a dummy cake. The results show that the participants prefer robots that make reasonable suggestions compared to random suggestions and that they perceive those robots to provide more creative support.

Keywords: Human-robot creative collaboration · Dialogue system · Cognitive architecture.

1 Introduction

Collaborative robots are being increasingly designed to help humans perform their tasks. For instance, robots are being designed to hand over tools in a way that makes it convenient for the humans to immediately use them [16] and optimize task allocation in collaborative assembly [23]. In this paper, we instead investigate intuitive and creative collaboration, motivated by potential applications in the arts, entertainment, and in assistive robotics where robots carry out creative tasks that users are not collocated to directly execute themselves. Creativity can help people lead happier and more meaningful lives [12] [4]. Given the importance of creativity, there has been a rise of interest in fostering human creativity through human-robot creative collaboration. Research in this area shows that interacting with a social robot can facilitate creativity. For instance, Kahn et al. show that participants are able to generate more ideas in a Zen rock garden task when a humanoid robot encourages them and shows relevant images and video clips compared to the base condition in which the participants access the same information through self-paced PowerPoint presentation [10]. Oliveira et al. show that children are able to generate more ideas during storytelling when playing with an intervention tool robot that exhibits creative behaviors compared to the control condition in which

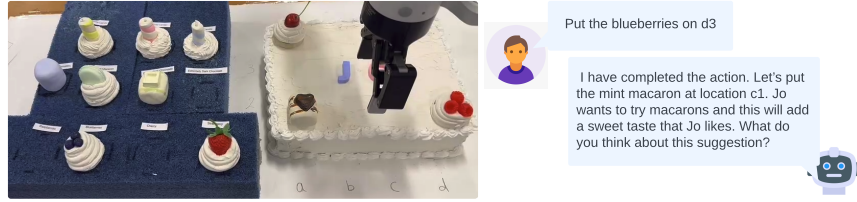


Fig. 1: The cake decoration task setup

the robot is turned off [3]. Buyukgoz et al. find that participants' generated more new ideas when interacting with a robot exhibiting medium level of proactive behavior in the form of verbal interruptions compared to when interacting with a robot exhibiting high level of proactive behavior [6]. However, not a lot of work has been done to investigate the effects of different modes of creative agent initiative in the form of suggestions and justifications using architectures that both support interactive task dialogues and ground the dialogues in physical manipulations. We present such an architecture and hypothesize that:

- H1: Task dialogues with a robot will increase task satisfaction and experience.
- H2: A robot that offers suggestions and task-specific reasoning behind suggestions is perceived as more intelligent.
- H3: A robot that offers suggestions and task-specific reasoning behind suggestions provides more creative support to the user.
- H4: Users prefer collaborating with a robot that offers task-specific reasoning behind its suggestions.
- H5: Users are less likely to reject suggestions from the robot that offers task-specific reasoning behind suggestions.
- H6: Users with higher creative self-efficacy are more likely to reject suggestions.

We present an interactive task dialogue system that interfaces with the Distributed Interactive Cognition Affect and Reflection Architecture (DIARC) to ground natural language descriptions of the task in the robot's actions [18]. The system is not only able to engage with the participant in a dialogue to complete a cake decoration task as shown in Fig. 1 but can also make placement suggestions for decorative items.

We validated this dialogue system in a within-subjects user evaluation in which participants worked with three different versions of the robot to complete the task. We collected participants' ratings of creative support, perceived robot intelligence, task experience and satisfaction as well as open-ended feedback on all three conditions. Finally, we compared the relative utility of the three robot configurations. The results showed that while the robot's participation in interactive task dialogues tended to increase the user's perceived creative support, the robot's ability to give task-specific justifications for suggestions had a large positive effect on not only the creative support the user received but also the overall user experience. In summary, our contributions are as follows:

- We present an architecture that integrates a large language model into a robot cognitive architecture to support interactive task dialogues.
- We present user evaluation results on a co-creative task with different modes of agent initiative.

2 Background

2.1 Human-AI Mixed-Initiative Co-creation

Mixed-initiative co-creation (MI-CC) refers to the process during which a human and an artificial intelligence (AI) system take the initiative to contribute to the creation of an end product [22]. Previous research has investigated mixed initiative co-creative systems that create digital products. For example, Alveraz et al. explore MI-CC in storytelling [2] and Liapis et al. explore it in game design [14]. We focus on architectures that collaborate during ideation through task dialogues while interactively carrying out the physical manipulations for the user. Our system is an instance of an MI-CC system that brings the creation into the physical world. Interaction with a physical MI-CC that communicates ideas through natural language and is solely responsible for the physical manipulations has different properties than previously-studied systems. Since users may perceive the system as more of a creative partner than a tool, it is important to directly investigate it as we do in this work.

2.2 Human-Robot Creative and Physical Collaboration

There has been extensive research on human-robot physical collaboration in which success is clearly defined. Popular tasks for evaluating human-robot physical collaboration include assembly tasks [1, 19], construction tasks [21], sorting tasks [9], human-robot navigation in search and rescue [8], and efficient tool handover [16]. In contrast, our work addresses creative tasks in which the goals of the tasks are loosely defined and humans have more freedom to complete the tasks according to their preferences. More specifically, our work addresses human-robot co-creation tasks during which both the human and the robot participate in the ideation process.

Recently, there has been a rise in interest in human-robot creative collaboration. However, research in this area has focused on tasks that either lack a physical aspect or on tasks in which the robot and human both take part in the physical aspect. For example, Sandoval et al. [17] investigate people’s perception of robot’s creativity through a storytelling game with a Wizard of Oz setup, Thorn et al. [20] explore observers’ perception of a robot after watching videos of it dancing to a human pianist’s music, Kahn et al. [10] measure participants’ increase in creativity when collaborating with a robot that encourages them to think of more ideas during a Zen rock garden task, and Buyukgoz et al. [6] study the effects of a robot’s proactive behavior in the form of verbal interruptions on participants’ creativity. None of these tasks requires the robot to physically manipulate objects. On the other hand, Lin et al. [15] present a co-creative process in which a robot needs to move a pen across a piece of paper to sketch with the participant and Law et al. [13] showcase a collaborative design task in which the robot and the human simultaneously manipulate a tangible user interface. While these tasks involve physical manipulations, both the human and the robot equally participate in the physical aspect. None of these tasks involve collaborative ideation through natural language dialogues. Our work addresses tasks in which collaboration through natural language happens during ideation and the physical manipulation is solely done by the robot. We expect that this approach will have applications in assistive robotics and service robotics.

3 Experimental System Overview

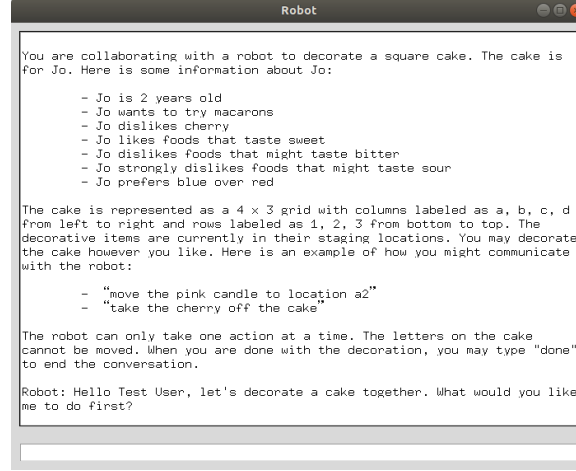


Fig. 2: The text user interface

To study the effects of interactive task dialogue on the human user’s task satisfaction and experience, we developed a system consisting of a dialogue manager that interfaces with DIARC capable of carrying out a cake decoration task [18].

The cake decoration setup consists of a textual user interface and a Gen3 Kinova robotic arm. The physical setup of the cake decoration is shown in Fig. 1. The cake is represented as a grid with columns labeled a, b, c, d and rows labeled 1, 2, 3. The robot is able to take two high-level actions: moving a decoration to a location on the cake or taking an object off the cake and putting it back in its staging location. The user communicates with the robot arm through a text user interface shown in Fig. 2.

3.1 Dialogue System

The dialogue system, based on GPT-4, generated both appropriate responses to human text inputs and made the appropriate robot action function calls based on the conversation. Fig. 3 illustrates the program flow of the dialogue system. After receiving human text input, the program appends it to the conversation history. The program dynamically fills in a prompt template with descriptions of the current state of the task. The system then queries GPT-4 with the filled-in prompt and the conversation history for what the response should be. If the output is “action”, the system chooses the appropriate action function signature and extracts parameter values from the conversation history. If the output is “suggestion”, the system either generates a “reasonable” suggestion or a random suggestion on what next action to take based on the experimental condition. If the output is “explain”, the system either generates a task-specific or a vague explanation based on the experimental condition. If the output is “other”, the system displays a pre-determined response to inform the user that it cannot respond to their request and redirects the user back to the cake decoration task. To ensure that the human and the robot take approximately equal initiative in the task, the robot alternates between prompting the human for what it should do next and giving a suggestion on what to

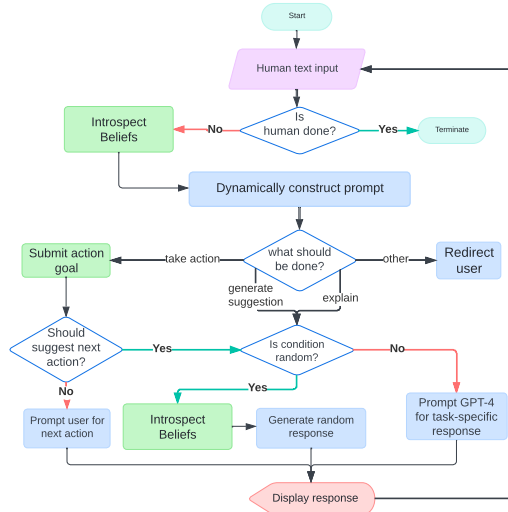


Fig. 3: Dialogue program flow

do next. More specifically, the robot generates a suggestion in two cases: when directly prompted by the human and when it has just successfully completed an action commanded by the human. The robot does not generate a suggestion when it has just completed an action suggested by itself. As a result, the robot proactively suggests a next move only about half of the time. The dialogue system interacts with DIARC when carrying out a robot action and when introspecting the robot’s beliefs about the task state. We use DIARC as the robot’s cognitive architecture because it is easily configurable. In our case, the LLM-based dialogue system acts as the natural language processing components—namely, the Natural Language Understanding (NLU), Natural Language Generation (NLG) and the dialogue manager components in the cognitive architecture. The robot has two user-facing actions: moving a decoration to a location on the cake and taking a decoration off the cake. We built “action scripts” for both actions using pre-existing primitive actions in DIARC. When trying to execute an action, the dialogue system submits a string consisting of an action script signature with parameter values, e.g., *TakeOffCake(cherry)* as an action goal to the goal manager. The dialogue system then receives an action status after the completion or failure of the action from the goal manager and reports it to the user. When constructing the task description prompt for the LLM, the robot receives the robot’s current beliefs represented as predicates that evaluate to true about the world, e.g., *at(cherry, a3)* and *freecakeloc(b1)*. The dialogue system translates these into natural language descriptions, e.g., *the cherry is at location a3* and *the location b1 on the cake is not occupied* and fills in a prompt template with these descriptions of the current state of the task. Similarly, when generating a random suggestion on what valid action to take next, the dialogue system queries DIARC for beliefs on what objects exist, i.e., *object(X, physobj)*, which objects the robot is able to pick up, i.e., *canpickup(X)*, which locations are unoccupied, i.e., *freecakeloc(X)*, and which objects are on the cake, i.e., *on(X, cake)*. The dialogue system generates a random valid suggestion by randomly choosing an object to place at a random unoccupied location or by randomly choosing an object that’s on the cake to be taken off. We show examples of conversations between the dialogue system and participants below.

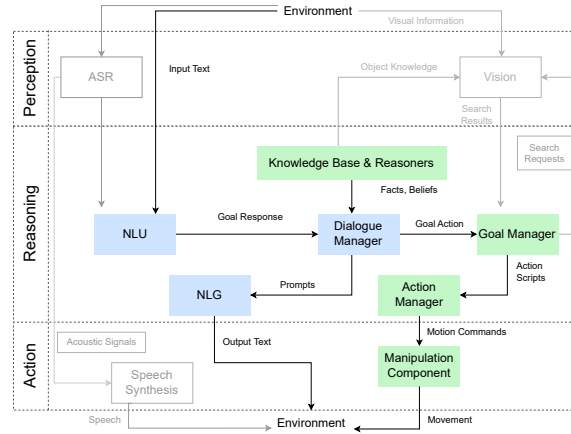


Fig. 4: The DIARC cognitive architecture with a goal manager, an action manager, and a customized dialogue manager, NLU and NLG components. The blue components were replaced by a Large Language Model (LLM).

3.2 Experiment Conditions

To study the effects of having versus not having dialogue and different modes of agent initiative, we implemented three experiment conditions: no dialogue, dialogue with random suggestions and agent initiative, and dialogue with reasonable suggestions and agent initiative. Participants work with a robot arm to decorate a cake for a customer named Jo with the following given information in all three conditions.

- Jo is 2 years old
- Jo wants to try macarons
- Jo dislikes cherry
- Jo likes foods that taste sweet
- Jo dislikes foods that might taste bitter
- Jo strongly dislikes foods that might taste sour
- Jo prefers blue over red

No Dialogue In the no dialogue condition, the robot serves as an assistive tool in picking up and placing the decorations. The robot follows commands and gives no suggestions on what to do next.

Robot interaction in the no-dialogue condition

PARTICIPANT: put the mint macaron at a2
 SYSTEM: The robot is processing your message...
 SYSTEM: The robot has successfully completed the action
 PARTICIPANT: take the cherry off the cake
 SYSTEM: The robot is processing your message...
 SYSTEM: The robot has successfully completed the action
 ...
 PARTICIPANT: done
 SYSTEM: Goodbye! Have a nice day!

As shown in the conversation above, the participant receives no feedback other than system messages reporting the robot’s status in the third-person perspective.

Random Suggestion In the random suggestion case, the robot serves a more active role in designing the cake by making suggestions. After action completion, the robot generates a response containing a random suggestion and a vague justification. The suggestion is randomly selected from the set of all valid next actions. The justification is randomly selected from a set of predefined sentences expressing the tautological idea that “this is a good action because it is good.”

Robot interaction in the random suggestion condition

ROBOT: Hello Human, let’s decorate a cake together. What would you like me to do first?

PARTICIPANT: take the raspberry off the cake

SYSTEM: The robot is processing your message...

ROBOT: I have completed the action. Let’s move the cherry to location c3. This decision is smart because it makes a positive difference. What do you think of this idea?

PARTICIPANT: ok. let’s do it!

SYSTEM: The robot is processing your message...

ROBOT: I have moved the cherry to location c3. What would you like me to do next?

PARTICIPANT: put blueberries on b1. This will provide a nice $y=x$ symmetry!

SYSTEM: The robot is processing your message...

ROBOT: I have completed the action. Let’s move the white chocolate cube to location d1. Doing this action is good because it’ll lead to better outcomes. What do you think of this idea?

PARTICIPANT: Nah! Suggest something else!

SYSTEM: The robot is processing your message...

ROBOT: Let’s move the blue marshmallow to location d3. This is wise, as it means things will improve. What do you think of this idea?

...

PARTICIPANT: done

ROBOT: Goodbye! Have a nice day!

As shown in the conversation above, the robot alternates between prompting the participant for the next action and suggesting the next action. The predefined justifications for suggestions are designed so that the length of the responses in the random condition are comparable to that of the responses in the reasonable condition, controlling for robot response length in the study.

Reasonable Suggestion In the reasonable suggestion case, GPT-4 is prompted to generate suggestions based on the task description and the robot’s current belief state of the task. GPT-4 generates suggestions in the following format: “Let’s {description of action}. {reason for selecting the action}. {ask what the human user thinks of this suggestion}” (the same format of response as that of the random suggestion robot).

Robot interaction in the reasonable suggestion condition

ROBOT: Hello Human, let’s decorate a cake together. What would you like me to do first?

PARTICIPANT: Put the blueberries on d3

SYSTEM: The robot is processing your message...

ROBOT: I have completed the action. Let's put the mint macaron at location c1. Jo wants to try macarons and this will add a sweet taste that Jo likes. What do you think about this suggestion?

PARTICIPANT: I agree, lets do it

ROBOT: I have successfully placed the mint macaron on the cake. What would you like me to do next?

...

PARTICIPANT: done

ROBOT: Goodbye! Have a nice day!

As a result, the robot gives task-specific suggestions based on Jo's preferences.

4 Study Methodology

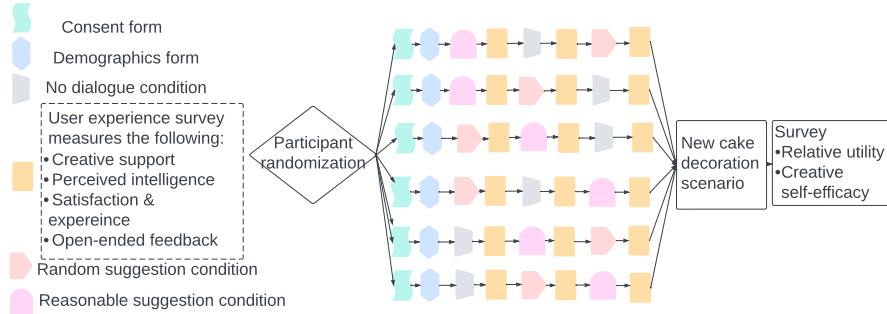


Fig. 5: Experiment Procedure. Participants were randomly assigned to one of the six orders of the three conditions. Participants started with filling the consent and demographic forms and ended with filling the creative self-efficacy survey. We applied the surveys using Qualtrics.

We conducted a within-subject user study during which participants experienced all three conditions mentioned above. Participants were told that they could decorate the cake however they liked. After each condition, the experiment facilitator manually reset the cake to the starting state shown in Fig. 1. As shown in Fig. 5, we randomized the order in which the participants experienced the three conditions. Furthermore, we performed block randomization to counterbalance the number of participants that were assigned to each order.

4.1 Measures

After each condition, participants completed a survey about their experience. The questionnaire measured participants' perceptions of the creative support received from the robot and the robot's intelligence, participants overall task satisfaction and experience, and open-ended feedback that asked the participant to "describe [their] experience working with the robot".



Fig. 6: New cake decoration scenario. Participants were asked to choose a robot to work with on the cake for a new customer

Suggestion Rejection Rate We analyzed the log files for the rate at which the participants accepted, rejected, and modified the robot's suggestions in both the random condition and the reasonable condition.

Creative Support Participants were asked to rate the following statements on a 5-point Likert scale. We designed the statements to measure the degree to which interaction with the robot supported artistic creativity. The items were based on the Creative Support Index, a validated measure for AI tools [7], modified to reflect the collaborative role of the robot and the physical nature of the task.

- Interacting with the robot stimulated my own thinking
- Interacting with the robot was helpful when decorating the cake
- Interacting with the robot was effective in completing the cake decoration task
- Interacting with the robot helped me come up with a better cake design
- Interacting with the robot led me to think of more ways of decorating the cake

Perceived Intelligence We measured perceived intelligence of the robot using the Godspeed IV Intelligence scale [5]:

- 1=incompetent, 5=competent
- 1=ignorant, 5=knowledgeable
- 1=responsible, 5=irresponsible
- 1=unintelligent, 5=intelligent
- 1=foolish, 5=sensible

Overall Satisfaction and Experience We measured the participant's overall satisfaction and experience with the following scale:

- How satisfied are you with the decorated cake (1 - Very unsatisfied, 5 - Very satisfied)
- How pleasant is your experience in working with this robot on the task (1 - Very unpleasant, 5 - very pleasant)

Relative Utility We compared the utility of the no dialogue robot, the random suggestion robot and the reasonable suggestion robot by presenting the participant with another cake decoration scenario as shown in Fig.6 and asking them to choose a robot partner. We physically showed the participant a larger cake, more decorative items, and information about a new customer with different preferences. We then asked the participant to select a robot to work with and describe their reason for selecting this robot.

Creative Self-Efficacy After the conclusion of all three conditions, the participant rated the following statements on 5-point Likert scale to measure creative self-efficacy [11]:

- I think I am a creative person
- My creativity is important to who I am
- I know I can efficiently solve even complicated problems
- I trust my creative abilities
- Compared to my friends, I am distinguished by my imagination and ingenuity
- Many times I have proven that I can cope with difficult situations
- Being a creative person is important to me
- I am sure I can deal with problems requiring creative thinking
- I am good at proposing original solutions to problems
- Creativity is an important part of me
- Ingenuity is a characteristic which is important to me

The creative self-efficacy scale was administered at the end of the study to avoid any bias that might arise from subjects focusing on creative aspects early on.

5 Results and Discussion

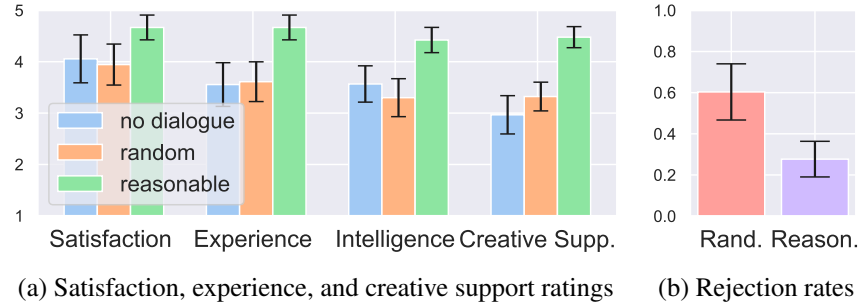


Fig. 7: User Satisfaction, Experience, Creative Support, and Suggestion Rejection Rates

We performed the Friedman’s test to analyze whether there was a significant difference in the ratings of task satisfaction, user experience, perceived intelligence, and creative support across the three conditions. We then performed the Conover’s post hoc comparisons to determine which conditions were significantly different from each other. We performed a chi-square test to analyze the suggestion rejection rate among the three conditions. Furthermore, we checked for normality and performed a linear regression analysis to determine the effect of creative self-efficacy on the suggestion rejection rate.

5.1 Participants

We recruited 18 university students. Of those participants, 9 identified as male and 9 female. Participants had an average age of 23, with a minimum of 18 and maximum of 32. Most participants (61.1%) have never worked with robots. 44.4% identified as being very familiar with technology; 38.9% identified as being moderately familiar; 16.7% identified as being not familiar at all. Participants had an average creative self-efficacy rating of 3.8 with a minimum of 2.7 and a maximum of 5.0.

5.2 Task Satisfaction and Experience

Table 1: Conover's Post Hoc Comparisons - Satisfaction

group1	group2	T-Stat.	df	W_i	W_j	p	P_{bonf}	P_{holm}
no dialogue	random	0.397	34	32.500	30.500	0.694	1.000	0.694
no dialogue	reasonable	2.480	34	32.500	45.000	0.018	0.055	0.037
random	reasonable	2.876	34	30.500	45.000	0.007	0.021	0.021

Table 2: Conover's Post Hoc Comparisons - Experience

group1	group2	T-Stat.	df	W_i	W_j	p	P_{bonf}	P_{holm}
no dialogue	random	0.463	34	27.500	30.000	0.646	1.000	0.646
no dialogue	reasonable	4.262	34	27.500	50.500	< .001	< .001	< .001
random	reasonable	3.799	34	30.000	50.500	< .001	0.002	0.001

Participants reported better satisfaction with the outcome in the reasonable condition compared to the random condition (all p values < .05). Participants also reported better experience in the reasonable condition compared to the random condition and the no dialogue condition (all p values < .001) as shown in Fig. (a), Table 1 and Table 2. However, the difference between the random condition and the no dialogue condition was not statistically significant. This result partially supports H1.

5.3 Intelligence

Table 3: Conover's Post Hoc Comparisons - Intelligence

group1	group2	T-Stat.	df	W_i	W_j	p	P_{bonf}	P_{holm}
no dialogue	random	0.087	34	29.000	28.500	0.932	1.000	0.932
no dialogue	reasonable	3.723	34	29.000	50.500	< .001	0.002	0.002
random	reasonable	3.809	34	28.500	50.500	< .001	0.002	0.002

While both the no dialogue robot and the random robot were rated as significantly less intelligent than the reasonable robot, we found no significant difference in intelligence ratings between the random robot and the no dialogue robot as demonstrated by Table 3. This result supports H2.

5.4 Creative Support

Table 4: Conover's Post Hoc Comparisons - Creative Support

group1	group2	T-Stat.	df	W_i	W_j	p	P_{bonf}	P_{holm}
no dialogue	random	1.679	34	22.500	32.500	0.102	0.307	0.102
no dialogue	reasonable	5.121	34	22.500	53.000	< .001	< .001	< .001
random	reasonable	3.442	34	32.500	53.000	0.002	0.005	0.003

As shown in Table 4 and Fig. (a), participants reported significantly more creative support from the reasonable robot than the random robot. Participants also reported significantly more creative support from the reasonable robot than the no dialogue robot. This finding aligns with H3, illustrating the utility of robot initiative in the form of suggestions and reasons in supporting creativity.

5.5 Suggestion Rejection

Table 5: Conover’s Post Hoc Comparisons - Rejection Rates

group1	group2	T-Stat.	df	W_i	W_j	p	p_{bonf}	p_{holm}
random	reasonable	2.222	17	31.500	22.500	0.040	0.040	0.040

We found that participants were significantly less likely to reject suggestions given by the reasonable robot compared to the random robot as shown in Fig. 7b and Table 5 ($p < .05$). This supports H5. We found no correlation between participants’ creative self-efficacy ratings and their suggestion rejection rates in either the random condition nor the reasonable condition. This does not support H6.

5.6 Relative Utility

Table 6: Multinomial Test for Relative Utility

	χ^2	df	p
Multinomial	30.333	2	< .001

Table 7: Descriptives for Relative Utility

condition	Percentage of Participants Expected: Multinomial	
base	0.000	0.333
random	0.056	0.333
reasonable	0.944	0.333

Participants overwhelmingly preferred to work with the reasonable robot on a similar task as shown in Table 6 and Table 7. 17 out of 18 (94%) participants preferred the reasonable robot whereas 1 out of 18 (<6%) preferred the random robot. None of the participants preferred the no dialogue robot. This supports H4. When reporting the reasons behind their choice, some participants highlight the reasonable robot’s effectiveness at reducing cognitive load by suggesting to place decorations that the participant “might have forgotten about” while others point to its ability to support creativity, commenting that they would be able to “combine their own creativity with the robot’s help”.

6 Discussion and Conclusion

The effects of interactive task dialogues in creative and collaborative settings underscores the substantial impact of task-specific justifications provided by a co-creative robot partner. Participants tend to give the highest ratings of support to the robot that incorporates task-specific justifications for suggestions into the dialogue. Results suggest that *task dialogue alone* is not enough to improve user satisfaction and experience; they showcase the impact of different modes of robot initiative, namely offering task-specific justifications while taking initiative significantly improves user experience. Participants perceive the reasonable robot as more intelligent and are less likely to reject its suggestions, indicating that perceived robot intelligence plays a significant role in participants’ willingness to accept the robot’s collaborative input. Furthermore, participants

overwhelmingly find the reasonable robot the most useful out of the three, further highlighting the importance of task-specific justifications in co-creative settings. Together these results show that the proposed architecture integrating LLM models to enable design suggestions through task-based dialogues does support users in making their own creative designs. It thus provides an important step towards developing robots that can aid people in carrying out creative tasks without limiting their creativity.

References

1. Admoni, H., Weng, T., Hayes, B., Scassellati, B.: Robot nonverbal behavior improves task performance in difficult collaborations. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 51–58. IEEE. <https://doi.org/10.1109/HRI.2016.7451733>
2. Alvarez, A., Font, J., Togelius, J.: Story designer: Towards a mixed-initiative tool to create narrative structures. <https://doi.org/10.48550/arXiv.2210.09294>, <http://arxiv.org/abs/2210.09294>
3. Alves-Oliveira, P., Arriaga, P., Cronin, M.A., Paiva, A.: Creativity encounters between children and robots. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. pp. 379–388. HRI '20, Association for Computing Machinery. <https://doi.org/10.1145/3319502.3374817>
4. Baer, J.: Content matters: Why nurturing creativity is so different in different domains. In: Beghetto, R.A., Sriraman, B. (eds.) *Creative Contradictions in Education: Cross Disciplinary Paradoxes and Perspectives*, pp. 129–140. Springer International Publishing. https://doi.org/10.1007/978-3-319-21924-0_8
5. Bartneck, C., Kuli, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots **1**(1), 71–81. <https://doi.org/10.1007/s12369-008-0001-3>
6. Buyukgoz, S., Pandey, A.K., Chamoux, M., Chetouani, M.: Exploring behavioral creativity of a proactive robot **8**. <https://doi.org/10.3389/frobt.2021.694177>, publisher: Frontiers
7. Cherry, E., Latulipe, C.: Quantifying the creativity support of digital tools through the creativity support index **21**(4), 21:1–21:25. <https://doi.org/10.1145/2617588>
8. Heintzman, L., Hashimoto, A., Abaid, N., Williams, R.K.: Anticipatory planning and dynamic lost person models for human-robot search and rescue. In: 2021 IEEE International Conference on Robotics and Automation (ICRA). pp. 8252–8258. <https://doi.org/10.1109/ICRA48506.2021.9562070>, ISSN: 2577-087X
9. Hinds, P.J., Roberts, T.L., Jones, H.: Whose job is it anyway? a study of human-robot interaction in a collaborative task **19**(1), 151–181. https://doi.org/10.1207/s15327051hci1901&2_7, place: US Publisher: Lawrence Erlbaum
10. Kahn, P.H., Kanda, T., Ishiguro, H., Gill, B.T., Shen, S., Ruckert, J.H., Gary, H.E.: Human creativity can be facilitated through interacting with a social robot. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 173–180. <https://doi.org/10.1109/HRI.2016.7451749>, ISSN: 2167-2148
11. Karwowski, M., Lebuda, I., Winiewska, E.: Measuring creative self-efficacy and creative personal identity **28**(1), 45–57, place: Republic of Korea Publisher: Korean Assn for Thinking Development
12. Kaufman, J.: Creativity as a stepping stone toward a brighter future **6**(2), 21. <https://doi.org/10.3390/jintelligence6020021>, number: 2 Publisher: Multidisciplinary Digital Publishing Institute

13. Law, M.V., Jeong, J., Kwatra, A., Jung, M.F., Hoffman, G.: Negotiating the creative space in human-robot collaborative design. In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. pp. 645–657. ACM. <https://doi.org/10.1145/3322276.3322343>
14. Liapis, A., Yannakakis, G.N., Alexopoulos, C., Lopes, P.: Can computers foster human users creativity? theory and praxis of mixed-initiative co-creativity <https://www.semanticscholar.org/paper/Can-Computers-Foster-Human-Users%E2%80%99-Creativity-Theory-Liapis-Yannakakis/a5cd2b0d1d4b8b99e96ce53be3786b7862740532>
15. Lin, Y., Guo, J., Chen, Y., Yao, C., Ying, F.: It is your turn: Collaborative ideation with a co-creative robot through sketch. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–14. CHI '20, Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376258>
16. Qin, M., Brawer, J., Scassellati, B.: Task-oriented robot-to-human handovers in collaborative tool-use tasks. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. pp. 1327–1333. <https://doi.org/10.1109/RO-MAN53752.2022.9900599>, ISSN: 1944-9437
17. Sandoval, E.B., Sosa, R., Cappuccio, M., Bednarz, T.: Humanrobot creative interactions: Exploring creativity in artificial agents using a storytelling game **9**
18. Scheutz, M., Williams, T., Krause, E., Oosterveld, B., Sarathy, V., Frasca, T.: An overview of the distributed integrated cognition affect and reflection DIARC architecture. In: Aldinhas Ferreira, M.I., Silva Sequeira, J., Ventura, R. (eds.) *Cognitive Architectures*, vol. 94, pp. 165–193. Springer International Publishing. https://doi.org/10.1007/978-3-319-97550-4_11, series Title: Intelligent Systems, Control and Automation: Science and Engineering
19. Shah, J., Wiken, J., Williams, B., Breazeal, C.: Improved human-robot team performance using chaski, a human-inspired plan execution system. In: *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. pp. 29–36. <https://doi.org/10.1145/1957656.1957668>, ISSN: 2167-2148
20. Thörn, O., Knudsen, P., Saffiotti, A.: Human-robot artistic co-creation: a study in improvised robot dance. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. pp. 845–850. <https://doi.org/10.1109/RO-MAN47096.2020.9223446>, ISSN: 1944-9437
21. Vasey, L., Nguyen, L., Grossman, T., Kerrick, H., Atherton, E., Thomasson, D., Cote, N., Research, A., Schwinn, T., Conti, M., Fitzmaurice, G., Menges, A.: Collaborative construction: Human and robot collaboration enabling the fabrication and assembly of a filament-wound structure. <https://doi.org/10.52842/conf.acadia.2016.184>
22. Yannakakis, G.N., Liapis, A., Alexopoulos, C.: Mixed-initiative co-creativity
23. Zhang, R., Lv, Q., Li, J., Bao, J., Liu, T., Liu, S.: A reinforcement learning method for human-robot collaboration in assembly tasks **73**, 102227. <https://doi.org/10.1016/j.rcim.2021.102227>