LARGE LANGUAGE MODELS IN THE DESIGN PROCESS: A WORKSHOP ON THE POSSIBILITIES OF HUMAN-AI COLLABORATION THROUGH KNOWLEDGE FORMATION

A. Mastroianni, L. Rampino, F. Figoli

Politecnico di Milano (ITALY)

Abstract

With the rise of Artificial Intelligence (AI), Large Language Models (LLMs) are proving to be a valuable technology for enhancing the designer's creativity in the design process (DP). For instance, LLMs can synthesize information, brainstorm ideas, and simulate user opinions. We analysed the implementation of LLMs in the early phases of the DP, the Discover and Define phases, to investigate AI's impact on the designed output and on the collaborative sphere of the design team. To achieve such objectives, after a comprehensive literature review and an initial test of the technology, we organised a workshop, in which 86 design engineering students took part. They were required to redesign a medium complexity product, using ChatGPT 3.5 throughout the early stages of the DP, sided by the "Prompt-chaining Cards" tool. We assessed the participants' initial knowledge, biases, and attitudes towards the LLM, and tracked the evolution of these perceptions throughout the workshop. Moreover, we conducted a qualitative analysis of the generated data, the research paths taken by the participants in collaboration with AI, and the identified needs and traits of the fictional users. The workshop's results enabled us to elucidate the opportunities and implications arising from the collaboration between product design students and LLMs. In particular, the results highlight the importance of "prompting" the correct inputs to the AI and the need for a critical discussion and analysis of the AI-generated data within the design team. Additionally, incorporating LLMs into a design team creates a set of specific challenges to be addressed, particularly a lack of communication between the human and machine counterparts, that ask for a high level of awareness of the process, a mastery of the technology and strategic management skills to effectively take advantage of the AI. In conclusion, our knowledge contribution to the discussed topic is an "integration framework" intended to help designers maximize the benefits of using LLMs in the first phases of the DP while highlighting the technology's current limitations.

Keywords: Artificial Intelligence, ChatGPT, Human-AI collaboration, Product Design, Design process, Prompt chaining cards.

1 INTRODUCTION

Our society is getting more and more complex, as we are witnessing significant shifts in our lives, ways of living, working and interacting with others and our surroundings. We are on the verge of the 5th Industrial Revolution (IR) [1], a period in which technology is converging towards collaboration rather than competing, as humans and machines are starting to engage with Artificial Intelligence (AI). The spread of AI has seen an extension to most fields of society, generating interest for its versatility, speed, optimization and wide-range capabilities [2],[3]. Among the most widespread AI models are Large Language Models (LLMs) [4], which play a pivotal role in AI diffusion, given their accessibility and ability to unlock new *interfaces* and communication forms. Such developments are bringing substantial changes and challenges in how designers approach *wicked problems* [5], fostering the emergence of new practices within the field of design. There is a growing concern that traditional design tools and skills might not be sufficient to fully encompass the scope of future project needs, and the presence of AI is emerging as a key candidate for future human-machine collaboration [7], marking the next steps in the evolution of the design field. This is further supported by the increasing body of academic research around the topic [7], [8]. Integrating AI in the design field is marking a transformative shift, with its ability to perform intricate cognitive functions [9] and interactions with humans through a reciprocal flow of inputs and outputs [10]. Researchers are now moving beyond the concept of augmentation and stimulation of human performances and creativity [11], considering AI as an active, collaborative partner during the creative workflow [7],[12]. However, introducing AI in a collaborative design environment, generates profound implications, both in operational and relational terms. From the literature review, three main factors emerge. The *performance* of the specific AI in use strongly influences the relationship with the technology, both practically and relationally, as poor AIs' performances generate a negative effect on the team ones [13]. *Trust dynamics* are at the essence of the interaction human-AI, given the

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)

influence of individual experiences, sensitivities, and emotions of each user. In collaborative environments, particularly those incorporating non-human actors, trust becomes vital [14], [15]. *Over-trust* in the technology might induce early complacency and misuse of AI, while *under-trust* could lead to underestimation of the resources, inefficient oversight, and biased task distribution [7]. Lastly, the lack of *visibility* of the thinking paths of AI [15], and the difficulty in understanding how to control it, challenges our notions of technology. The determination of who - human or AI - when and how has control during the interactions is evidenced as a fundamental step to achieve a correct collaboration and consequently enhancing the DP [16],[17]. Even though researchers are expanding our view in the sphere of AI, essential knowledge regarding the impact of LLMs in the design process (DP) is still underexplored [7]. Unlike traditional tools, LLMs can actively contribute to developing novel insights and solutions by leveraging their large datasets and algorithms [11]. Similar to human creativity, LLMs can retrieve and recombine existing knowledge, making it appear novel. These capabilities position LLMs as potential creative entities, able to mimic cognitive processes and generate innovative recombination of ideas. However, LLMs generate outputs in response to input in the form of written *prompts*. Thus, the creative work of an LLM is strictly dependent on the human input [11], [18]. Moreover, while LLMs can generate content based on specific inputs, they cannot independently initiate the need for ideation. This means that, although LLMs can effectively support the creative process and produce ideas [10], [19], it is ultimately up to humans to interpret these ideas. For that reason, it is crucial to recognize the complexities of the deployment of LLMs for practical applications, preventing misuse and ensuring benefits during the different phases of the DP.

Building on the current state-of-the-art, an exploratory study was conducted to investigate both LLM's impact on the outputs' quality of the *discover* and *define* phases - the first two stages of the *double diamond design process* (DP) by the Design Council - and the collaboration between human-AI within design teams. The study involved students from the course *Design theory and practice* - held by one of the authors at Politecnico di Milano - who were asked to share their perspectives on incorporating ChatGPT during the initial phases of the DP. Additionally, the study sought to examine the potential impact of introducing a specific tool designed to help students face LLMs. This paper focuses on a specific part of the master's degree thesis of the paper's first author.

2 METHODOLOGY

The study was carried out from September to December 2023, incorporating two main experimental stages with two different scopes: Auto-ethnography and Collaborative Scenario.

The first stage of the qualitative research was conducted on an individual basis - auto-ethnography -, and focused on a preliminary assessment and evaluation of LLM's capabilities This analysis was fundamental to build, test and define prompt strategies and structures, necessary to craft a tool for the following phases of the workshop, the prompt chaining cards. The second stage of the research, on which this paper focuses, was a collaborative scenario conducted in the form of a workshop in the Design & Engineering Master's at Politecnico di Milano, involving 86 students as participants. This phase focused on evaluating the efficacy of a collaborative scenario in which humans and the ChatGPT 3.5 LLM worked synchronously within the DP, in a 4-hour design activity. The workshop simulated a condensed DP: the discover, and define phases were performed in collaboration with ChatGPT, and a short redesign activity without AI. The 86 participants were divided into 16 groups composed of 5 to 6 students, as they had to collaborate during the different phases of the workshop to redesign a medium-complexity home appliance. To collect the generated data, we had direct access to each group workflow and interactions with the LLM, through the online platform Figma. Furthermore, to capture students' perceptions and thoughts on the topic, we collected qualitative observations of teams' behaviours during the workshop, and designed two questionnaires, one before and one after the workshop, to have a deeper understanding of how their personal experiences and opinions shifted during the design activity. In the two questionnaires, we aimed to understand participants' initial knowledge and usage of LLMs, and their attitude towards it before and after the workshop experience. We focused on grasping their general acceptance of LLMs within the early stages of the DP, the trust in the generated suggestions and the perceived influence that AI had on their decision-making. We collected data on participants' awareness and sensitivities, identifying nuanced interaction patterns, perceptions and potential biases towards the usage and collaboration with LLMs. A note has to be made on the crafted tool designed for the collaborative scenario, the Prompt chaining cards. Based on the insights of the auto-ethnography, the cards intended to provide practical guidance to inexperienced users navigating the complex communication processes with LLMs. Developed around two communication approaches, the Dialogical and Constrained one, each card was purposely crafted to focus on a single DP activity, allowing students to decide which one was coherent with their objectives and to draw inspiration from their structure. The tool serves as an exploratory guideline, offering a flexible path for approaching the design process (DP) in collaboration with LLM technology rather than a framework to follow rigidly.

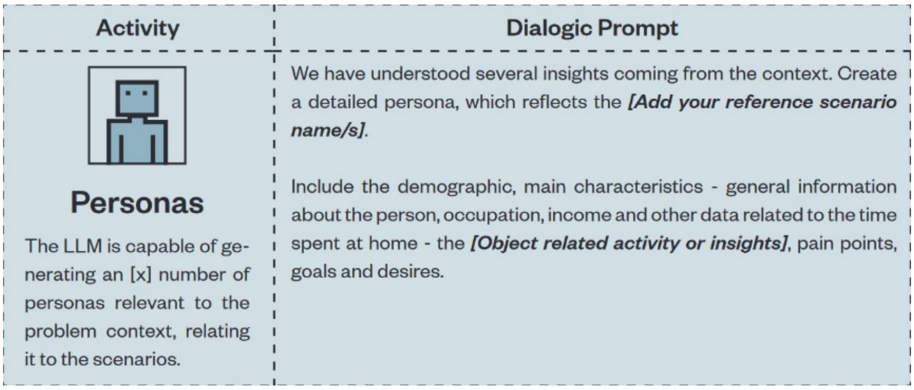

Figure 1. Structure of the "Personas" prompt chaining card.

3 RESULTS

Participants responded to the first questionnaire, aimed at gauging their general attitudes and biases toward using LLMs in the DP. The responses revealed an overall lack of common understanding about AI, resulting in a wide array of answers. Few students demonstrated a solid understanding of AI. Others had a dystopian view, driven by a sentiment of fear coming from the current AI's capabilities and their future potential. However, the diversity and breadth of the answers underline that participants lack a general idea and comprehension of what AI is, as they had their perceptions, views and beliefs on the matter. Notably, almost all participants (94%) have already used Chat GPT 3.5, and 71% have tried to use LLMs either in the *discover* or *define* phase. These initial questions proved that the immediate diffusion of LLMs is sided by a strong interest in using and understanding such technologies even in the design field, showing a general welcoming attitude in using LLMs in the DP. The question in *Fig.2* intended to understand the students' level of trust and acceptance towards the LLM's generated outputs. The initial participants' answers were compared with those from the final questionnaire at the end of the workshop activities. Firstly, participants were quite hesitant to trust AI outputs, with responses generally indicating neutral stances that leaned slightly towards a positive and cautious trust.

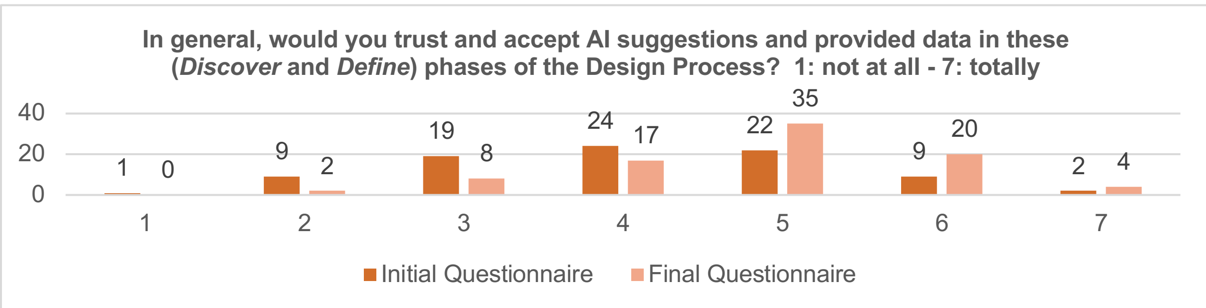

Figure 2. Participants' trust and acceptance answers.

However, the comparison highlights a significant inclination towards change in trust perception, as more than half of the participants - 52 - showed a reassessment and change of their answer. The majority of such changes were directed towards a more favourable position, with 39 participants changing their judgement to a higher trust. Such shifts highlight that students were susceptible to such experiences, rapidly improving their perception of the matter. Overall, although most participants showed increased trust, a comparison of each participant's responses with those of their teammates revealed that a few groups experienced trust in the LLM in a fragmented manner, with some groups perceiving the reliability of the outputs differently. Furthermore, these differences within the same team underline that some participants had a *lonely* interaction and relationship with the LLM. In the activity, a few groups had some difficulties in sharing information between the team members, and as a result, each one was iterating autonomously with the LLM, without sharing the data within the team and, at the same time, magnifying

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)
This version should not be distributed since it may change prior to final publication

the quantity of outputs to analyse. Such behaviour might conceal a sense of disconnection among the team's participants. The last question (*Fig.3*) focused on understanding the perceived influence that the LLM had in the decision-making processes during the activity. The initial results show that the LLM is perceived as an impacting technology within its implementation. Already before taking part in the workshop activities, the majority of the participants believed that AI would have a significant impact on the team's decisions.

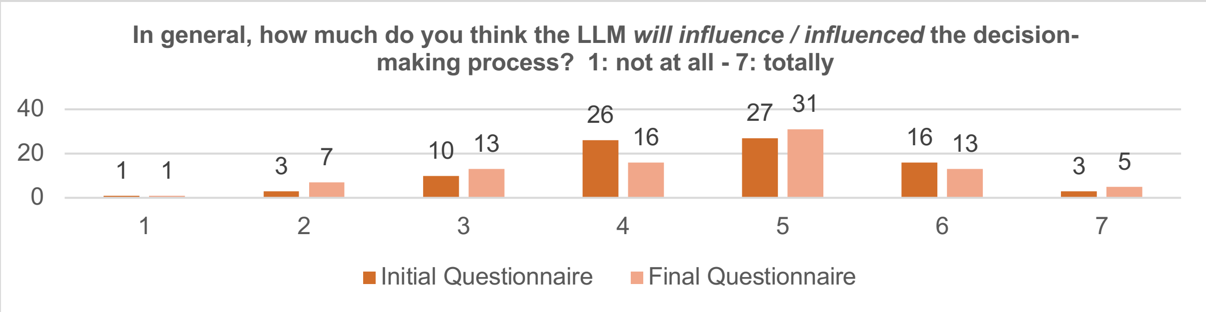

Figure 3. LLM's decision-making influence.

By comparing the results with the final questionnaire, it emerges that few participants changed their opinion. 25 % of participants ended the workshop believing that AI had limited influence on their decision-making, while 56% had the opposite perception. These results, both before and after the activity, show that most students had a predisposition to evaluate the LLM as a strong voice in the decision-making process. An interesting finding is that, by observing the significant increase in *trust*, the same rise did not happen with the perceived influence in the *decision-making* processes. Relating the single participants' answers in *Fig.2* and *Fig.3*, it emerges that the participants' relation between *trust* and *decision-making* is not linear. The participants who felt less influenced by AI still had a neutral or major trust in the LLM's answers. Moreover, cross-linking the answers it appeared that a wide percentage of the participants with the highest levels of *trust* towards the LLM, were the ones that felt more controlled by it. Even if students experienced the same activity, they responded subjectively, leading to different perceptions regarding using the LLM in the early stages of the DP. The groups that evidenced balanced levels of *trust* and *influence* in decision-making, were able to *discuss* and *plan* their objectives ahead of each iteration with the AI, subsequently taking advantage of the *prompt-chaining cards* and shaping their communication towards the LLM to better suit their purposes. That particular relationship underlines critical factors regarding the *awareness* of using AI. The results evidenced that a strong sentiment of *trust* towards AI's capabilities might lead to overexposure to its suggestions. It is important that users retain their role of *decision-makers*, as underlined by one participant of *Group 9*: "*I think AI can be very useful to enhance these phases with proper information and time optimization, but the person who is using AI needs to filter through all that information as well.*"

The complex balance already evidenced by the questionnaires was analysed through an extensive examination of each groups' processes and interactions with the LLM and within the team. The first stage of the DP, the *discover* phase, appeared to be a moment of extensive usage and deployment of AI, as we can observe in *Fig.4*.

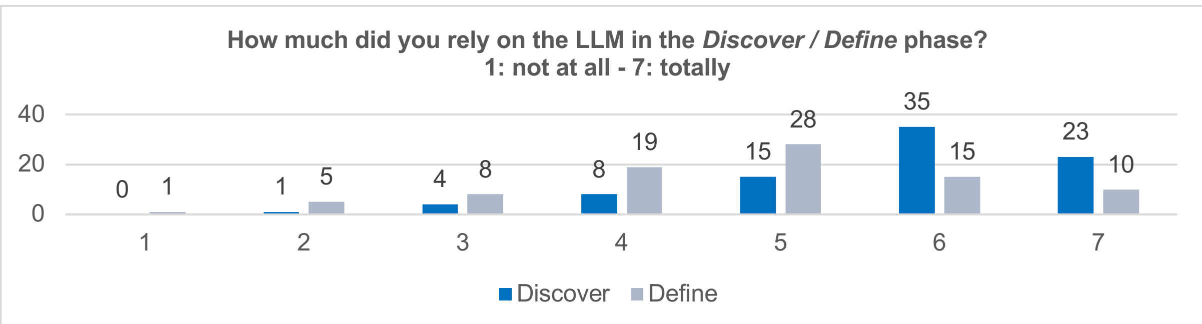

Figure 4. LLM's reliance in each phase.

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)
This version should not be distributed since it may change prior to final publication

However, participants relied slightly less on AI during the *define* phase, allowing more room for personal reflection and work. Still, participants appreciated the organising skills offered by the LLM, as conversing with the AI model and asking it to organise the obtained data, was perceived as a *validation activity* by the students, to confirm their assumptions through the LLM's outputs. A student of *Group 2 stated*: "*In the Define phase, ChatGPT was a little less helpful, as it was very good at finding information rather than developing solutions, I believe. For this reason, it was mainly used for the mapping part as a verification for the process.*"

Looking back to the initial questionnaires and confronting the responses to the results of *Fig.4*, the LLM was widely associated with a process of expansion of knowledge, such as acquiring, collecting and enriching the team's vision and data towards a specific topic, particularly an unexplored one (namely the re-design of a new product). A student from *Group 7* underlined that: "*AI is a strong tool to boost efficiency and obtain information that would be hard to find or understand otherwise. It is very powerful as an assistant or external source; however, it shouldn't be the only source of information, as it may become an oracle if not filtered properly.*" This inclination and predisposition withhold the prior perceptions of the students and could be explicable through their inclination to trust the AI as a form of "*wider*" and "*superior knowledge*", sometimes seen as an "*oracle.*" As identified also in the *auto-ethnography* prior to the workshop, it is observed that *prompt iteration* helps the LLM to better grasp the desired outputs of the team, improving the overall communication - as a *Group 8* student evidenced: "*The LLM started working with more precision as soon as it we got deeper with the research. Different iterations make the algorithm more mature to release linked insights to our proposed route, as it created connections with different fields, considering our point of view.*" Some groups asked the LLM to *simulate* specific *brainstorming scenarios* or *design awards presentations* to enhance their vision and gather insights from new perspectives. On the contrary, some groups decided to generate *unstructured prompts* and conversations with the LLM, without following the *prompt chaining cards*, and misaligning the order of the activities of the DP (*Fig.5*).

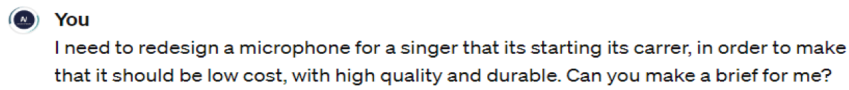

Figure 5. Group 11 - first prompt iteration.

As a result, the LLM generated shallow and general responses, leading to incorrect iterations within the team and igniting a *trust misalignment* towards the LLM that impacted the team's decisions and performances. This initial behaviour is both relatable to an initial *over-trust* of the participants in the capabilities of the LLM, and both to a *knowledge gap*, evidenced by an unclear comprehension of the *role* that the AI could play in their collaboration and by the poor *structuration* of the prompt. Nevertheless, we have to note that the groups able to sustain a coherent and iterative conversation with the LLM, using the *prompt-chaining* technique - and as such, incorporating the obtained insights into the following iterations - obtained valuable data in the following outputs of the LLM, which proposed different ways to frame the teams' identified problems and novel perspectives on the subject (*Fig.6*). As a result, the LLM generated shallow and general responses, leading to incorrect iterations within the team and igniting a *trust misalignment* towards the LLM that impacted the team's decisions and performances. This initial behaviour is both relatable to an initial *over-trust* of the participants in the capabilities of the LLM, and both to a *knowledge gap*, evidenced by an unclear comprehension of the *role* that the AI could play in their collaboration and by the poor *structuration* of the prompt. Nevertheless, we have to note that the groups able to sustain a coherent and iterative conversation with the LLM, using the *prompt-chaining* technique - and as such, incorporating the obtained insights into the following iterations - obtained valuable data in the following outputs of the LLM, which proposed different ways to frame the teams' identified problems and novel perspectives on the subject.

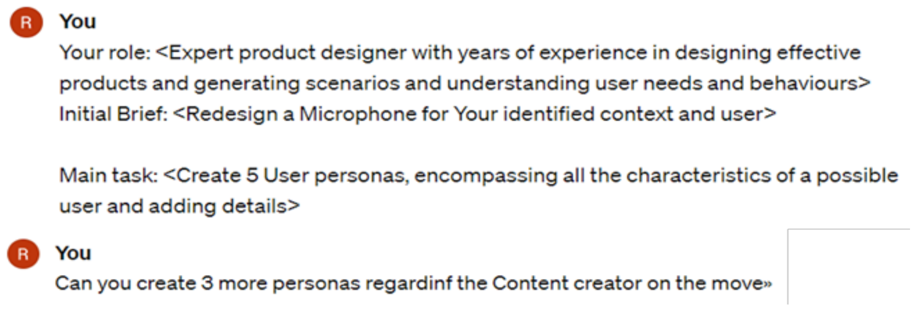

Figure 6. Group 10 - generating a chaining between two prompts.

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)
This version should not be distributed since it may change prior to final publication

From the analysis of the groups' interactions with AI, it clearly appeared that, for a correct balance in the collaboration human-LLM, crafting correct prompts was just part of the recipe. The teams needed to define their trajectory before conversing with the LLM, defining the *task* and the *role* that AI needed to cover in the activity. Taking the example of *Group 2*, they faced the activity with a step-by-step process of communication and review of the outputs of the LLM. After a preliminary use of the given *prompt-chaining cards* to explore the contexts of use for *portable lamps*, they fixed their objective of *simulating a contest* for people who desired this kind of object, in order to swiftly obtain multiple routes to explore and assess, consequently crafting the necessary prompts to bring the LLM to *personify* in this category of users. Such *awareness* of the process, maintained by some groups throughout the DP, let the LLM produce consistent outputs and enhanced the collaborative efforts, both with the LLM and within the team. The groups that were able to maintain a critical view of the process were also able to change their approach and adapt their *communication* with the LLM, steering towards their tasks and objectives over time, obtaining the most valuable information and outputs from it, as stated by *Group 2* student: "*Rather than saying that AI can influence the decision-making process, I can say that it helps to better analyse a wide range of users and scenarios. Moreover, it can provide very good insights into the different problems related to what you want to design. Of course, it is up to the designer to analyse and filter the individual pieces of information and especially to understand which ones can lead to the desired result.*"

The last phase of the workshop, the *redesign*, conducted by the students without the LLM, consisted of generating a concept starting from the insights obtained during the previous phases, as we had to link the obtained results and observations to the final process and sets of data that each group used as a basis for their work. Generally speaking, almost all of the groups showed sufficient to a high level of consistency from the collaboration with the LLM within the DP, as the presented results showed good levels in terms of quality, variance and innovation. The groups that were able to take advantage of communication and prompting to expand their views towards their objectives, maintaining a *balanced collaboration* with AI, showed interesting results that distinguished in terms of quality of the identified needs, variety and effectiveness of the process.

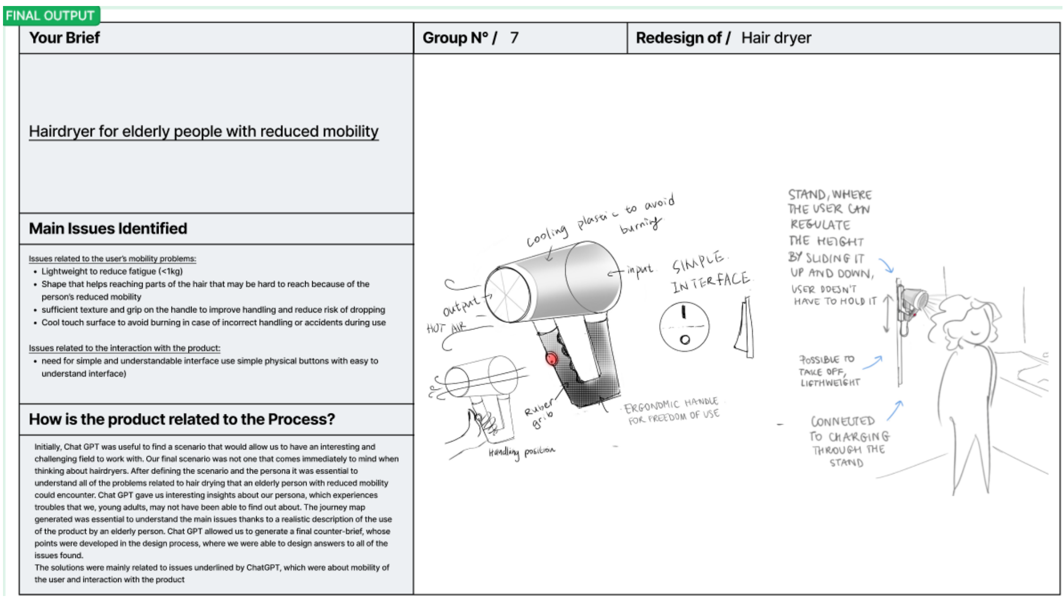

Figure 7. Group 7 final output. Redesign of a hair dryer.

As we can see in *Fig.7*, some results explored new routes that changed perspectives on the object to redesign, both in terms of *user groups* and in terms of *contexts of use*. Such hidden perspectives, as underlined by the same groups, were driven by a coherent and fruitful interaction between the team and the LLM, that led to empathising with different scenarios and categories of users - such as elderly people for *Group 7*. The obtained results did not evidence criticalities regarding lack of variety or qualitative wrong insights. However, some groups, particularly the ones redesigning the *portable lamp*, showed some repetitions regarding context and user categories. Similarly, some insights reported by the AI, although being utilised correctly by some groups, had to be further explored and verified more critically, involving specific use scenarios that immediately questioned the feasibility of the proposed path - e.g. a *hair dryer* for outdoor enthusiasts. Lastly, we can observe that the results of the groups that were unable to achieve correct *communication* could be regarded as insufficient. Such outputs did not focus on a specific route,

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)
This version should not be distributed since it may change prior to final publication

insight or user, and consequently, did end up with shallow outputs, evidencing the importance of a proper process and a balanced collaboration. In *Fig.8* we can observe the output of *Group 12*, coming from a process of *unstructured prompts* and *iterations* with the LLM. The identified insights lacked depth, focusing on technicalities that went beyond the purpose of the activity, remaining shallow in the understanding of the user and context. Moreover, the data used in the final output was directly transcribed from the AI answers, lacking a *reassessment* of the LLM's outputs and *awareness* of their position.

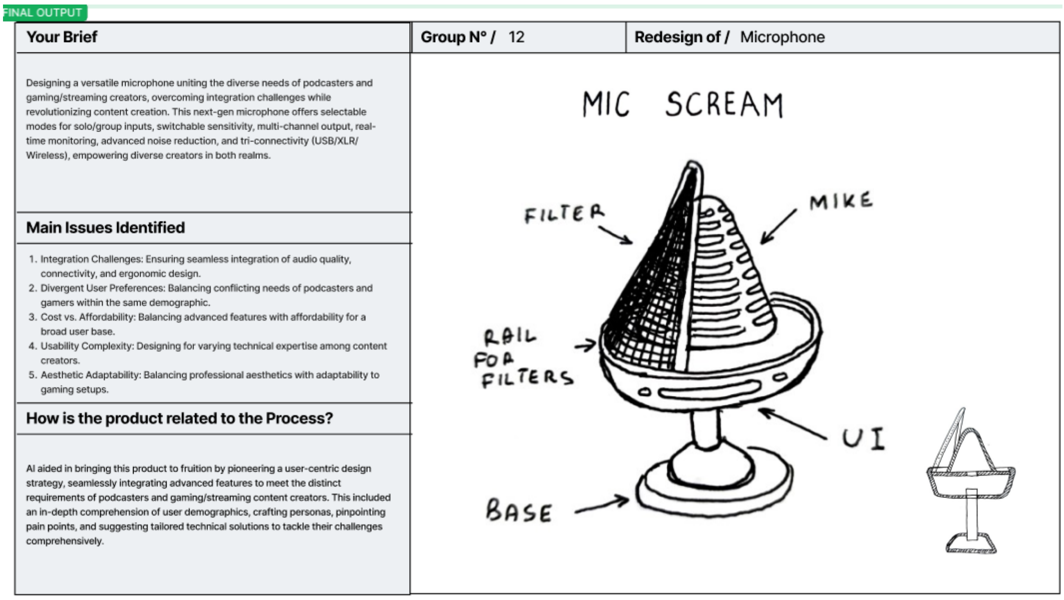

Figure 8. Group 12 final output. Redesign of a microphone.

4 DISCUSSION

Several considerations can be made when describing the results. From the questionnaires, students showed an overall welcoming attitude and curiosity towards LLMs, particularly regarding their interest in using them in the early stages of the DP. This positive attitude was tempered by subsequent questions regarding trust in the outputs of the LLM and the perceived highly influential role of AI in decision-making processes. However, this neutral and positive position at the beginning of the workshop swiftly changed after the activity, as we have seen in *Fig.2* and *Fig.3.* Such changes emphasised a certain level of influenceability, that can be mainly explained by two factors.

Knowledge gained during the workshop: participants have been able to take advantage of their acquired knowledge, particularly regarding the communication towards the AI through the *prompt-chaining cards*, leading to effective usage of the LLM within the DP. This improved the quality of the LLM's generated answers and, consequently, directly increased the participants' *trust* and *acceptance* of the LLM's suggestions during the workshop. *Lack of knowledge prior to the workshop*: many students showed a lack of knowledge both in the theoretical and practical notions about such technology.

The increasing *acceptance* and *trust* likely emerge from the participants' limited prior exposure and experience of LLMs, leaving them without a firm opinion. Furthermore, the complex relationship between some groups and AI, marked by the divergent opinions within teams, underscored the apparent detachment some students experienced from their peers. As perceived by *Group 1*: "*The AI was very useful in the early stages and helped enrich our early perspective. We had some problems with the amount of information that held us from discussing the data between us."* This sense of detachment arose not only from miscommunication with the AI but also from internal team dynamics that evolved over time, leading to a growing disconnection both from the LLM and among teammates. When new technology is approached without sufficient knowledge and with excessive trust in the LLM, participants can become overly reliant on the machine. This reliance can lead to the AI overshadowing the team dynamics, as students may not fully understand their own roles versus the role of the LLM. Therefore, early experiences in collaborating with AI systems are important for shaping students' understanding of how LLMs might function within the design process. From the analysis of the results, we can summarise *two critical factors* in facilitating effective human-AI collaboration: *communication* and *awareness*. Effective communication, both with the structuration of prompts for the LLM - *external* - and for an assessment of the data within the design team - *internal* - emerged as a critical factor in eliciting valuable insights and fostering an appropriate level of *trust*. At the same time, *awareness* of LLM's roles, capabilities and limitations as well as the designer's role in the process, were fundamental to achieve a fruitful collaboration.

One of the main challenges encountered with this new technology was the *knowledge gap*, prior to the workshop, in managing communication interactions, often characterised by an initial overestimation of the LLM's capabilities. Users sometimes assumed that the LLM could interpret and respond to human language as adeptly as a fellow human. This led to an expectation that the LLM would unconditionally understand prompts as intended. However, the analysis of the results revealed that this notion of mutual understanding was frequently challenged by the students' initial experiences with the outputs. *Incorrect prompting* techniques and inadequate communication approaches often led to *unsatisfactory interactions* with the LLM. As observed, the output of each LLM mainly depends on its input. Comprehending the fundamental role of *crafting* appropriate *prompts* for different tasks was the first, crucial step to obtaining deep and appropriate responses from the LLM, as most students achieved sufficient levels of *external communication* through the usage of the *prompt-chaining cards*. Meanwhile, using *wrong prompt techniques* and crafting unbalanced requests to the AI led some groups to improper LLM usage, which generated shallow and low-quality insights, negatively affecting the subsequent decisions of the team. In this scenario, a snowball effect of *misalignment* between humans and AI can begin, potentially hindering human-LLM collaboration to the point of breakdown. Consequently, the second factor that influenced the *collaboration* human-AI was the *internal communication* that is built within the group itself. One of the most notable observations from the workshop was the initial struggle among group members to evenly distribute the use of the LLM within the team. This challenge also affected their ability to effectively share and communicate insights. Some teams did not collectively assess the LLM's outputs, which led to a lack of critical discussion about the validated insights and highlighted how some students were overly influenced by the AI. This communication breakdown revealed an *over-reliance* on AI and *over-trust* in their *prompting* skills. As a result, data was evaluated less critically, leading to insufficient team *realignment* between each step and iteration, with the evaluation process becoming superficial, failing to question the validity, variability, or depth of the LLM's responses. The absence of *internal communication* diminished the quality of the outputs, both for an absence of insight sharing and discussion and for a low level of iterations with the LLM. Consequently, participants tended to perceive the LLM as an unquestionable *oracle.* The *communication framing scheme* in *Fig.9*, highlights the identified players with the arrows pointing at the identified path of iterations of the communicative process, emphasizing the highlighted factors. Moreover, it evidences the relation between the *LLM-generated* knowledge and the *human-generated* one.

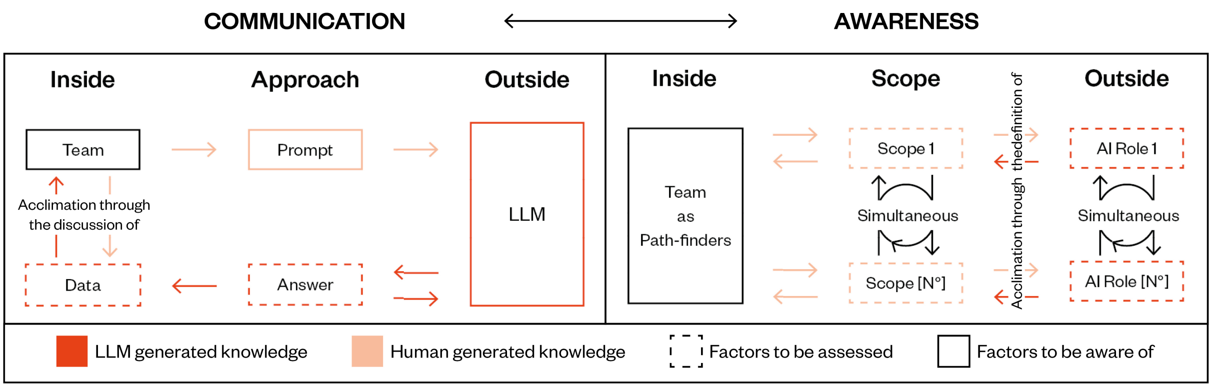

Figure 9. Communication and Awareness framing schemes.

The concept of *communication* is strictly related to the participants' *awareness* of the collaborative endeavour in which groups are involved. During the workshop, some participants developed their own vision and *awareness* of their relation with the LLM, comprehending and discerning the different possibilities that build up within that collaborative environment and that consequently modify the direction of their process, shaping the functions, roles and contributions of both the designer and AI. The results showed that there is a delicate balance between human intuitions and AI guidance, with decision-making processes shaped by several factors: each group's communication skills, their vision of the project trajectory, their interaction with the AI, and the roles assigned to both the team and the LLM. As synthesised in the *awareness framing scheme* in *Fig.9*, for each specific *scope* that had to be achieved, there was a specific path and *role* that better suited the purpose, as the versatility and flexibility

This article will be included in the ICERI2024 Proceedings (ISBN: 978-84-09-63010-3)
It will be fully citable as soon as it appears in IATED Digital Library (library.iated.org)
This version should not be distributed since it may change prior to final publication

of the inputs that were crafted made it possible to adjust the process of the LLM towards virtually infinite routes. Moreover, the scheme evidences the concept of having *simultaneous scopes* in our collaboration, which will lead to the corresponding *simultaneous roles* of AI - the LLM could both simulate an interview for a certain persona and subsequently generate new questions for a different one, being both a user and a teammate in the process. Being conscious and *aware* of the different features of the technology, guide the users to assume specific behaviours for the specific desired outcomes. Conversely, insufficient *awareness* may result in a *passive utilisation* of AI, with users inadvertently becoming instruments for the machine itself. As it emerged from the results, students were encouraged to develop their vision and deepen their current understanding of LLMs. This aspect is vital, as it shows the capability to increase the students' *awareness* through design experiences, highlighting the value of training as an effective solution to address human-AI integration.

Therefore, the moments of initial *acclimation* with the LLM and the correct understanding of the functions that the AI has to assume during the collaboration became a fundamental part of its integration within the DP. Being *aware* of the LLM behaviours and criticalities positioned the user at the centre of the collaboration. Designers had to adapt and change their *communication*, to achieve the best results and not fall into the criticalities of the LLM, thus being *aware* of their path, *approach* and the *roles* of each actor. We can define the designer as a *Pathfinder*, as a figure with a good knowledge of Large Language Models - as well as the Design Process - comprehending their limitations and possibilities, capable of effectively *communicating* with the machine, optimising and constantly evaluating the path to follow, *adapting* the communication towards the most suitable *approach* needed to achieve the defined tasks. This scenario synthesises the findings of the research, thus encapsulating the concept of *aware communication* into a figure that possesses profound skills for communicating with the AI while possessing a deep knowledge of the technology, being able to assess the collaboration objectives, remaining efficient, conscious and sensible during all the different steps of the iteration, constantly *evaluating* the *outcomes* of the machine to direct the collaboration towards the identified path.

Our *integration framework* positions the designer as a *pathfinder* in the human-AI collaborative process. This framework, in *Fig.10*, emphasises the need for designers to form knowledge in two directions, the *communicative* area to overcome the language barrier and master how to speak with AI, and a comprehensive understanding of the current possibilities of the technologies, to be *aware* of the different roles that the LLM can cover, adapting the communication to achieve the desired results. Such a process of adaptation, synthesised in the concept of *aware communication*, can lead the user to effectively collaborate with the LLM, maximising the benefits while mitigating potential pitfalls, such as *over-reliance* or *under-utilisation* of AI.

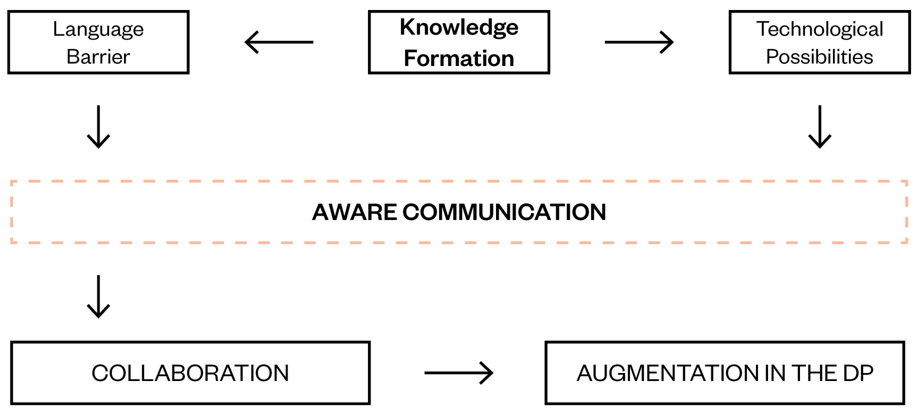

Figure 10. LLM integration framework.

5 CONCLUSION

The integration of Large Language Models in the early stages of the design process presents both significant opportunities and challenges. The study revealed a generally positive welcoming attitude towards LLMs, coupled with a high degree of influenceability in the students' perceptions of AI's capabilities, underscoring the importance of providing structured, comprehensible, and personalisable introductions to these technologies in design settings. Our findings highlight two critical factors affecting the success of human-AI collaboration: *communication* and *awareness*. Effective communication, both with the LLM - *external* - and within the design team - *internal* - proved to be crucial for generating valuable insights and maintaining a balanced level of *trust*. At the same time, *awareness* of LLM's roles, capabilities and limitations as well as the designer's role in the process, were fundamental for achieving a fruitful collaboration. Students who developed a nuanced understanding of their relationship with the AI were able to leverage its capabilities while retaining their critical role as *pathfinders*.

Future research should focus on developing more complete training tools for designers, creating collaborative LLMs, which could minimise the downsides of current human-AI interactions and exploring the impact of LLMs integration in a full-length DP.

REFERENCES

[1] S. M. Noble, M. Mende, D. Grewal, and A. Parasuraman, "The Fifth Industrial Revolution: How Harmonious Human–Machine Collaboration is Triggering a Retail and Service [R]evolution," Journal of Retailing, vol. 98, no. 2, pp. 199–208, Jun. 2022, doi: 10.1016/j.jretai.2022.04.003.

[2] M. Chui, E. Hazan, R. Roberts, A. Singla, K. Smaje, and A. Sukharevsky, The economic potential of generative AI The next productivity frontier. McKinsey & Company, 2023.

[3] S. Makridakis, "The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms," *Futures*, vol. 90, pp. 46–60, Jun. 2017, doi: 10.1016/j.futures.2017.03.006.

[4] S. Huang and P. Grady, "Generative AI's Act Two," Sequoia Capital. [Online]. Available: https://www.sequoiacap.com/article/generative-ai-act-two/

[5] C. W. Churchman, "Guest Editorial: Wicked Problems," Management Science, vol. 14, no. 4, pp. B141–B142, 1967.

[6] Rampino L. Evolving Perspectives in Product Design: From Mass Production to Social Awareness, 2018 (Franco Angeli, Milano).

[7] Figoli F. A., Mattioli F. and Rampino L. Artificial intelligence in the design process: The Impact on Creativity and Team Collaboration. 2022 (Franco Angeli, Milano).

[8] F. Isgrò, S. D. Ferraris, and S. Colombo, "AI-Enabled Design Tools: Current Trends and Future Possibilities," in *[ ] With Design: Reinventing Design Modes*, G. Bruyns and H. Wei, Eds., Singapore: Springer Nature Singapore, 2022, pp. 2836–2847. doi: 10.1007/978-981-19-4472-7_183.

[9] G. Martini, C. Cautela, and G. Carella, "LLMs Facing Up to Human Creativity," MA thesis, Politecnico di Milano, 2023.

[10] J. Haase and P. H. P. Hanel, "Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity," *Journal of Creativity*, vol. 33, no. 3, p. 100066, Dec. 2023, doi: 10.1016/j.yjoc.2023.100066.

[11] Z. Zeng, X. Sun, and X. Liao, "Artificial Intelligence Augments Design Creativity: A Typeface Family Design Experiment," in *Design, User Experience, and Usability. User Experience in Advanced Technological Environments*, vol. 11584, A. Marcus and W. Wang, Eds., in Lecture Notes in Computer Science, vol. 11584., Cham: Springer International Publishing, 2019, pp. 400–411. doi: 10.1007/978-3-030-23541-3_29.

[12] G. N. Yannakakis, A. Liapis, and C. Alexopoulos, "Mixed-initiative co-creativity," 2014, [Online]. Available: https://www.um.edu.mt/library/oar//handle/123456789/29459

[13] G. Zhang, A. Raina, J. Cagan, and C. McComb, "A cautionary tale about the impact of AI on human design teams," *Design Studies*, vol. 72, p. 100990, Jan. 2021, doi: 10.1016/j.destud.2021.100990.

[14] E. J. de Visser *et al.*, "Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 459–478, May 2020, doi: 10.1007/s12369-019-00596-x.

[15] J. Liao, P. Hansen, and C. Chai, "A framework of artificial intelligence augmented design support," *Human–Computer Interaction*, vol. 35, no. 5–6, pp. 511–544, Nov. 2020, doi: 10.1080/07370024.2020.1733576.

[16] S. Amershi et al., "Guidelines for Human-AI Interaction," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland Uk: ACM, May 2019, pp. 1–13. doi: 10.1145/3290605.3300233.

[17] J. D. Zamfirescu-Pereira, R. Y. Wong, B. Hartmann, and Q. Yang, “Why Johnny Can’t Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–21. doi: 10.1145/3544548.3581388.

[18] H. Zhang, C. Wu, J. Xie, Y. Lyu, J. Cai, and J. M. Carroll, “Redefining Qualitative Analysis in the AI Era: Utilizing ChatGPT for Efficient Thematic Analysis,” Sep. 19, 2023, *arXiv*: arXiv:2309.10771. Accessed: Nov. 25, 2023. [Online]. Available: http://arxiv.org/abs/2309.10771

[19] A. B. Kocaballi, “Conversational AI-Powered Design: ChatGPT as Designer, User, and Product,” 2023.