



Design process augmentation

# Amplifying the propositional learner using LLMs

---

Design theory unification

---

A holistic view on design process augmentation

---

An investigation of LLM creativity

---

An augmentation system interface

---

Tom Kortenbach

Master Thesis

Strategic Product Design

# Abstract

Current literature that aims to describe the use of (Gen)AI in the domain of design is hindered by the lack of conceptual integration of design theories which in turn blurs the connection between the design process and human-AI interaction and collaboration frameworks. This thesis' objective is to contribute to the conceptual understanding of LLM-based design process augmentation. This thesis shows how such LLM-based design process augmentation can look like by establishing 1) a time-based framework of a human – Augmentation-System collaborative design process, 2) a perspective on the design process augmentation capabilities of LLMs, and 3) an augmentation system architecture and interface for the practical implementation of these theoretical considerations. Other contributions that have enabled these three main contributions include 1) the unification of existing design theories (problem-solution co-evolution, situated Function-Behavior-Structure (FBS) framework, Concept-Knowledge theory, and Uncertainty Driven Action (UDA) model) to attend to a wider range of person-related characteristics required for the description of a collaborative process of design, 2) the connection between human intellect augmentation and creativity literature, resulting in 3) the application of creativity literature in the domain of GenAI and specifically LLMs, and 4) the application of the layers of behavior as described in the UDA (Uncertainty Driven Action) model for the creation of a computational augmentation system. Finally, a reflection on the feasibility and desirability of different forms and applications of augmentation systems is provided. The results of this thesis have created a fundament for further research into the direction of the initial objective, as the results have increased the conceptual power to describe and explain process, output, and other augmentation-related phenomena connected to the domain of human-LLM (or GenAI) co-design.

**Author**

Tom Kortenbach

**Supervisory team**

Rebecca Price, chair

Alessandro Bozzon, mentor

**Delft University of Technology**

Industrial Design Engineering

**Date**

5th of November 2024

# Preface

Long ago, I set my mind on becoming a car designer. I started my own hypothetical car company, Tobach (a contraction of my first name and surname), and started to create many, many drawings. This fascination for cars has led me to the faculty where great car designers such as Adriaan van Hooydonk (BMW) and Laurens van den Acker (Renault) were born and the faculty where I have enjoyed many years of educational and social prosperity: Industrial Design Engineering (IDE) at the TU Delft.

It has been during this IDE journey, and specifically during the Strategic Product Design (SPD) master program, that I have started to build a fascination for Artificial Intelligence (AI) and its capabilities relative to that of human intelligence. The existence of a computational alternative to specific forms of human intelligence forces you to reflect on how you work yourself. It forces you to look at the way in which you process information and use it to create new things. It forces you to think about how it is even possible that we humans can *design*. Even more important, if we have figured out a way in which a form of AI can design, do we want it to replace our human design practice? Driven by these fundamental questions about our human purpose, both in the biological as well as economic sense, this thesis was born<sup>1</sup>.

---

<sup>1</sup> The original graduation project brief can be found in appendix A



## Acknowledgements

I would like to express my gratitude to my thesis supervisors, Rebecca Price and Alessandro Bozzon, for their incredibly meaningful perspectives and feedback over the course of the graduation project. Within our often short but effective meetings we have been able to get to the core issues or to new perspectives on the existing content that have not only kept this thesis on its rails, but also led to meaningful academic contributions. Besides, I would like to thank Milene Gonçalves and Tilman Dingler for their participation in the focus group and their perspectives on topics such as creativity theory, design theory, and human-AI interaction literature. Finally, this thesis would not have been up to this quality level without the prior research into LLM creativity under the supervision of Pan Wang, since a large part of my perspective on doing research, my writing, and my knowledge in the domains of LLMs and creativity theory have been gained during that period.

# Content

<b>Content.....</b>	<b>6</b>
<b>1. Introduction .....</b>	<b>9</b>
1.1 Motivation .....	9
1.2 Scope, relevance & objective.....	10
1.4 Limitations of existing research .....	11
1.5 Method & thesis outline.....	13
1.6 Contributions.....	14
<b>2. Research methodology .....</b>	<b>15</b>
2.1 The implications of prior research for the direction of this thesis.....	15
2.2 A thesis that emerged from an iteratively prototyped process .....	15
2.3 Ensuring internal validity.....	17
2.4 Ensuring practical relevance .....	17
2.5 A note on the technological developments of LLMs .....	17
<b>3. Defining LLM-based designer augmentation .....</b>	<b>19</b>
<b>4. On the integration of design theory and human-AI collaboration frameworks.....</b>	<b>21</b>
4.1 Approach .....	21
4.2 Human-AI collaboration .....	21
4.3 Design theory .....	25
4.4 Towards a unified design process framework for augmentation .....	27
4.5 Towards the integration of AI-based systems in design .....	28
<b>5. A unified design process framework .....</b>	<b>30</b>
5.1 Approach .....	30
5.2 Design as co-evolution of problems and solutions.....	32
5.3 Cognitive co-evolution model.....	32
5.4 The origin & evolution of theories about problem-solution co-evolution.....	33
5.5 Idea representations: Function, Structure & Behavior (FBS).....	35
5.6 Ideas, concepts, and knowledge: Concept-Knowledge design theory .....	41
5.7 Schemata as knowledge representations .....	42
5.8 Instantiated schemata, Situated FBS framework & C-K theory .....	42
5.9 ‘Problems’ and ‘solutions’ expressed in a unified design process framework .....	45
5.10 The relation between and progression of understanding and knowledge .....	48

5.11 Towards interaction points between designer and augmentation system .....	51
5.12 The relation between types of goals and design actions .....	55
5.13 Relating the layers of behavior to the unified design process framework.....	56
5.14 From design actions to mental iteration stages .....	58
5.15 Types and scales of mental iteration .....	64
5.16 A definition of the human actor in the design process .....	66
<b>6. Design process augmentation .....</b>	<b>69</b>
6.1 Approach .....	69
6.2 Defining design process augmentation at the level of mental iterations.....	69
6.3 An outcome-based definition of a successfully augmented process.....	73
6.4 Towards process-related attributes of augmentation .....	73
6.5 Novelty augmentation .....	76
6.6 Quality augmentation .....	81
6.7 Embedding the augmentation forms in a collaborative process .....	84
6.8 Conceptual framework for human-AS co-design .....	87
<b>7. Design augmentation capabilities of LLMs .....</b>	<b>90</b>
7.1 Approach .....	90
7.2 A conceptual understanding of how Large Language Models (LLMs) work .....	91
7.3 Theoretical approach: LLM augmentation novelty, quality, and iteration.....	95
7.4 A definition of LLM-based (& human) behavior .....	113
7.5 Practical approach: LLM augmentation novelty, quality & iteration .....	116
<b>8. An interface for an LLM-based design process augmentation system .....</b>	<b>122</b>
8.1 Approach .....	122
8.2 Augmentation system (AS) architecture .....	123
8.3 User Interface (UI) .....	126
<b>9. Socio-economic implications .....</b>	<b>137</b>
9.1 Persona-flow patterns as an organizational operating model .....	137
9.2 The role of humans and augmentation systems.....	138
9.3 Implications for organizational structures and data.....	138
9.4 Feasibility: data maturity .....	139
9.5 Desirability.....	140
<b>10. Discussion .....</b>	<b>142</b>
10.1 The research approach: putting the design process front and center .....	142

10.2 Theoretical contributions & limitations .....	143
<b>11. Conclusion.....</b>	<b>149</b>
<b>12. Reflecting on the (strategic) design practice.....</b>	<b>151</b>
12.1 Design propositions: to learn for yourself or for others? .....	151
12.2 From design in theory to design in practice .....	152
12.3 The future of Strategic (product) Design.....	154
<b>13. Bibliography .....</b>	<b>159</b>
<b>Appendices .....</b>	<b>167</b>
Appendix A: Graduation project proposal.....	167
Appendix B: Images of whiteboard drawings during literature review .....	173
Appendix C: Augmentation system interface design iterations .....	184

# 1. Introduction

## 1.1 Motivation

Since the introduction of ChatGPT<sup>2</sup>, Large Language Models (LLMs), that operate based on the transformer architecture (Vaswani et al., 2023), have claimed center stage in discussions around the nature of human intelligence and human productivity. LLMs are able to generate natural language to create text that is indistinguishable from that written by humans. Given the pursuit of the computational replication of human level intelligence in the field of computer science, there have been various attempts to create LLM-based agents, such as AgentGPT<sup>3</sup>, or to build multi-agent frameworks for LLM-based agent collaboration (as in (Liu et al., 2023) and (Wu et al., 2023)) that could autonomously plan and execute tasks based on a user-provided set of objectives. Given the generative capabilities of LLMs in the agent frameworks such as AgentGPT, LLMs are not dependent on human initiation in every step of the process. Instead of having passive computational systems that inform the human recipient on demand, these information systems can become an active participant in decision-making and the progression of processes. This presents both opportunities for the augmentation at the scale of individuals, but also brings forward a critical question about the balance of agency between human and artificial agents in making decisions.

There are two ways to look at these agents. From the perspective of an individual, one could argue that these agents augment the individual as they are extending the set of capabilities of the individual. However, from the perspective of an organization, these agents have the potential to automate certain tasks and, as such, to replace specific forms of human labor. This automation introduces several consequences. First, it introduces a tension between the efficiency incentives as embedded in the capitalistic economic system and the continuity of the societal stability, since the economic value of certain forms of human labor will change<sup>4</sup>. Moreover, automated methods can be misleading due to incorrect assumptions or biased training data (Shneiderman, 2020), and they may optimize fixed objectives that fail to adapt to changing conditions or adversarial manipulation (Heer, 2019). Besides, users of such systems may become overly dependent on computational recommendations, potentially resulting in a decline in critical thinking and loss of domain expertise (Heer, 2019). Although humans have biases too, formed by prior life events and other incentives originating from external factors, it seems a more resilient solution to have humans in control of the realization of decisions, as we can learn and be corrected. Most importantly, humans are social creatures that pursue some sort of social connection, ensuring at least a bare minimum of care for the future of our species.

---

<sup>2</sup> <https://chatgpt.com/auth/login>

<sup>3</sup> <https://agentgpt.reworkd.ai/nl>

<sup>4</sup> As reported in a study from Ernst & Young (EY): [Link](#) to report

Since design is a form of intelligent behavior (Smithers et al., 1990), this discussion about augmentation versus automation is highly relevant for the domain of design research. With regards to the embedding of AI in the design discipline, current design literature mostly builds on the concept of ‘hybrid intelligence’ (Dellermann, Ebel, et al., 2019) that is grounded in the rationale that humans and computers have complementary capabilities that, when combined, can augment one another. For example, Thoring et al. (2023) present various ways in which AI can augment the design process, and Berni et al., 2024 have constructed a framework of AI-support in design ideation including ethical considerations about bias, accountability, and transparency. Although they provide an initial perspective on augmentation in the domain of design, their work points at several big gaps in our understanding of how such a construct can be realized in practice. Thoring et al. (2023) present a research agenda containing a set of questions that mostly rest on a deeper question about how to create a GenAI-based system that both aligns with the desired direction of the designer and generates novel and quality suggestions in an iterative manner. Additionally, Berni et al. (2024) state that *“the presumed AI capabilities of disclosing links between stimulation means and designers’ actions [...] have not been exploited and turned into desired design applications”* (p. 1944).

## 1.2 Scope, relevance & objective

Given this background, this thesis is specifically interested in using LLMs to augment the design process using LLMs. Augmenting a designer with LLMs means amplifying the capabilities of a designer in conceiving the unknown based on the capabilities of an LLM-based system that interacts with the designer via natural language. The objective is to contribute to the conceptual understanding of design process augmentation by defining 1) how humans design, 2) what types of interactions lead to the augmentation of the design process, and 3) how an LLM-based augmentation system interface can realize the desired forms of augmentation. Below, the rationale behind this scope is further substantiated.

Both the nature of LLMs as well as the new relation they introduce between humans and information are important reasons to investigate the implementation of LLMs in the design process. First, the nature of LLMs is a good fit with the nature of the design process, as the process of design is fundamentally driven by the perception of incomplete knowledge and the resulting epistemic uncertainty (Cash et al., 2023, Cash & Kreye, 2017), and LLMs form a conversational and comprehensive knowledge base (Dwivedi et al., 2023). Second, LLMs present a new interaction between humans and information. Whereas previously one would search through static accounts of written language to find various information components that accumulated to the point where one could synthesize towards the answer of one’s question, LLMs take away some of that cognitive strain. Through a prompt, a written or spoken description of a question or command, LLMs are capable of directly addressing the question by formatting the information stored in the model such that it resembles the way a human would answer

a question. As such, the presentation of information is dynamic and customized based on the input prompt. This type of interaction between demand for information and the retrieval and presentation of that information presents new ways in which humans can access information. Furthermore, LLMs, being a form of generative AI (GenAI), can be a first step towards understanding how to use other forms of GenAI for design process augmentation, such as image-generating Generative Adversarial Networks (GANs) or forms of transformer-based applications. Although the image-generating capabilities are relevant for the field of design, this thesis focusses on LLMs given their language-interface that plays a central role in the future of human-computer communication and interaction.

#### 1.4 Limitations of existing research

Existing research provides an explorative perspective on the value of LLMs for the design process. First, existing research has already identified early signs of design-augmenting capabilities of LLMs. For example, Chen et al. (2024) state that LLMs are a powerful tool to support novice and experienced designers and foster innovative designs. Moreover, research that analyzed the effects of LLM-augmented morphological analysis revealed that LLMs are capable of improving the quality of the design outcomes by aiding the designer during the problem decomposition, idea generation, and idea combination phases (Chen et al., 2024). Besides, building on existing ‘supermind’ design methodology<sup>5</sup>, the ‘Supermind Ideator’ is proposed as a generative AI tool based on an LLM to support individuals in creative problem solving (Rick et al., 2023). This tool provides so-called *design moves* to boost an individual’s reflection and inspiration with the aim to reduce design fixation (Rick et al., 2023).

Although the examples of LLM-augmented design discussed above provide evidence of the positive potential of LLMs in the design process, there is a lack of connection between these empirical observations and existing literature about design process augmentation including its limitations (such as Thoring et al. (2023) and Berni et al. (2024)). This leads to a lack of conceptual descriptive power to explain both the successful as well the less successful LLM-based interactions with the design process. As identified by, and addressed in the remainder of this thesis, this disconnection can be attributed to four literature related issues, of which three are related to design theory.

1. First, human-AI collaboration frameworks (Döppner et al., 2019; Fabri et al., 2023; Rezwana & Maher, 2023; Viros-i-Martin & Selva, 2021) describe a range of person-related cognitive characteristics, but no single design theory covers all these cognitive aspects, whereas the theories together do almost cover all aspects. This observation aligns well with Berni et al. (2024) who state that there

---

<sup>5</sup> A [design methodology](#) developed by the MIT Center for Collective Intelligence

is still a scattered understanding of the potential effects of AI on the designer's creativity. Moreover, as identified by Thoring et al. (2023), GenAI's (and as such LLM's) *"potentials for and impact on the designer and the design process are yet unclear"* (p. 3352), as it can become yet another tool or replace some of the design aspects currently reserved for the human cognition.

2. Second, there is a lack of connection between the major design theories. This finding aligns well with Viros-i-Martin and Selva (2021) who argue that the progression of design theory research is hindered by the lack of a shared vocabulary and ontology. As such, if one would like to combine the different design theories to collectively cover the person-related characteristics, one would face a scattered vocabulary hindering the integration.

Both findings point at the need to unite existing design theories to be able to express a process of design in terms of the factors identified relevant in human-AI collaboration. This unified design process then forms the basis for defining how an external LLM-based augmentation system can interact with that process to augment it.

3. Third, if such a unified design process framework is conceived including a definition of how to augment it, there is no holistic theoretical view on the capabilities of LLMs to realize such augmentation practices. This leads to the observation from Berni et al. (2024) that there is a disconnection between conceptual accounts of design process augmentation and the development of practical design augmentation tools, in this case based on LLMs.
4. Fourth, even if design theories are unified and if an account of LLMs' augmentation capabilities is established, there is currently no augmentation system interface that could serve as a boundary object and testing ground for both the validity of the unified framework (Berni et al., 2024) and reflective practice on the desirability GenAI-based human intellect augmentation. For example, the LLM-based augmentation tool "Supermind Ideator" (Rick et al., 2023) lacks a connection to sophisticated accounts of design theory. Similarly, Allixr<sup>6</sup>, a tool designed to help designers design by enabling the creating of flow schema's consisting of several different generative AI models, lacks a theoretical fundament for explaining its effect on the design process. The result is that the effect of existing design augmentation tools on the designer's design progression can currently not be described by a conceptual framework and hence not be reasoned about in advance.

Given the aim to work towards both an internally and externally valid conceptual framework, it seems fitting to at least address the lack of an interface for an LLM-

---

<sup>6</sup> Allixr is the result of a master thesis created by Ruben Dekker (2022), that can be found [here](#). The Allixr tool can be found via the following [link](#).



augmented design process by translating the conceptual framework into a low-fidelity prototype interface. Future research can then use the interface concept to perform further testing.

### 1.5 Method & thesis outline

A literature review is used as a method to identify and connect existing accounts of research around the various topics introduced above, since the nature of the research gaps mostly surrounds a lack of unification and connection between existing accounts of research. Besides, experts from the domain of creativity theory and human-AI interaction have been consulted in this theoretical investigation to validate the findings and conclusions and contribute to the discussion part. Further explanations about the research methodology are presented in section two. The central research question that guides the investigation is as follows:

*How to augment the design process using LLMs?*

This question is subdivided into six questions that all attend to a different aspect of the main question. These six questions are defined as follows:

1. *What are the aspects that drive and steer the design process in a collaborative environment?*
2. *How should the design process be conceived such that it allows for the identification of the types of augmentation, and the points of interaction that manifest these augmentation practices?*
3. *How can the design process be augmented?*
4. *What are the capabilities of LLMs to realize design process augmentation practices?*
5. *How can design process augmentation practices be manifested in practice?*
6. *What are the socio-economic implications of LLM-based augmentation systems?*

The structure of this thesis is as follows. First, the methodology section (section two) outlines the way in which these research questions have been established and approached. Section three establishes a theoretical definition of design process augmentation, based on Engelbart's (1962) theory of human intellect augmentation that acts as the basis for the rest of this thesis. Section four attends to the first research questions and provides the theoretical frame for the remainder of this thesis. It captures the theories of design and frameworks of human-AI collaboration in design and addresses the four gaps in existing literature critical to the ability to describe the phenomenon of interest: LLM-based design process augmentation. Building on the first gap, and attending to the second research question, section five provides a detailed discussion of the main design theories to work towards a unified conception of the design process that covers a set of actor-related aspects necessary to describe the process of design in the domain of collaboration and augmentation. Section six uses this

design process framework to describe how the design process can be augmented, as such attending to the third research question. Attending to the fourth research question, section seven investigates the ability of LLMs to realize the set of design process augmentation practices formulated in section six. To answer the fifth research question, section eight designs a low-fidelity, non-functional prototype of an augmentation system and interface that materializes the descriptions of design process augmentation in a boundary object that can stimulate further discussion and reflection. Attending to the sixth research question, section nine discusses the different potential applications of augmentation systems, their feasibility and desirability, and the socio-economic consequences. Section ten provides a discussion on the value of the results of this thesis for the theoretical landscape and identifies threads for future research. Finally, after the conclusion in section eleven, section twelve translates the theoretical and conceptual findings to insights that are relevant for both current Strategic Product Design (SPD) students, and future form of the (strategic) design practice.

#### **Takeaway boxes**

For those readers who want to quickly scan the content of this thesis, each section contains a (set of) green-colored 'takeaway box(es)' that present(s) a more easily digestible take on the content within each section.

### **1.6 Contributions**

This thesis mainly contributes to the field of design theory, while also presenting insights relevant to the fields of research that aim to further our understanding of human-AI collaboration and interaction. The contributions of this thesis include a(n)

- unified time-based design process framework based on existing design theories including problem-solution co-evolution, situated Function-Behavior-Structure (FBS) framework, Concept-Knowledge theory, and Uncertainty Driven Action (UDA) model.
- design process augmentation framework describing the six ways in which the design process can be augmented.
- time-based human-GenAI co-design process framework based on the unified design process framework and design process augmentation framework.
- theoretical perspective on the design process augmentation capabilities of the current generation of Large Language Models (LLMs) (up to OpenAI's GPT-4o).
- LLM-based Augmentation System architecture and interface.
- perspective on the socio-economic implications of different types of augmentation systems based on their feasibility and desirability.
- perspective on the nature, place, and value of (strategic) design in organizations and the implications for the future of the SPD master.

## 2. Research methodology

---

The literature review, as executed in this thesis, has been explorative rather than planned. Instead of crafting a plan for execution, the execution itself has iteratively crafted the plan. In other words, the process of conceiving the thesis has itself been one of iterative design and prototyping.

---

### **2.1 The implications of prior research for the direction of this thesis**

The starting point for this thesis has been shaped by prior investigations in the domain of the capabilities of LLMs. During another literature review into the creativity of LLMs, my objective has been to provide a perspective on the capabilities of LLMs from the perspective of creativity. This research led to a more intuitive conception of the LLM as a ‘creator’ within a certain ‘environment’. This then simplified the comparison between human and LLM capabilities to arrive at a view on the complementary capabilities of both types of actors.

This research had two effects on the conception of this thesis. First, it provided a basic understanding of the LLM’s capabilities in the context of this research, since design and creativity can be researched through the lens of each other (Taura & Nagai, 2010), and therefore allowed for the exploration of two unknown domains – both the fields of human-AI collaboration, augmentation, and interaction, as well as the field of design theory – leading to the integration of GenAI-based systems and the design process. Second, it inspired the selection of literature addressed in this thesis. For example, based on my ever-growing understanding of the design process, an intuitive conception of design process augmentation started to emerge. This conception grew towards the point where a clear similarity between design process augmentation and definitions of creativity arose. This led to the connection between design process augmentation and creativity theory literature. As such, creativity theory literature is discussed in several sections of this thesis, forming both the theoretical basis for the explanation of the processes that lead to a successfully augmented design outcome, as well as an investigation in the design process augmentation capabilities of LLMs.

### **2.2 A thesis that emerged from an iteratively prototyped process**

This thesis uses an iterative way of executing a literature review. The reason for this approach is the fact that this thesis aims to address a construct that is grounded in multiple research domains with each domain containing a gap that is related to the

subject of this thesis. An increased understanding in one domain thus leads to a potential reformulation of the thesis' investigation or the identified problem in another domain and vice versa. As such, this thesis has followed the approach of picking one domain as a starting point to then identify other relevant domains and gaps, and to go back and forth between the different domains of literature to evaluate on and update the direction of the investigation. Figure 1 provides a visual overview of the four research domains that form the basis of this thesis, including the six iterative learning loops. These six loops describe the process of going back and forth between existing and new knowledge within every domain to both identify the relevance of the new knowledge and the completeness of the existing knowledge.

Building on the previous knowledge in the domain of creativity literature and the capabilities of LLMs, this thesis' iterative process of literature investigation started within the domains of design theory and human-AI collaboration and augmentation literature. After many cycles of going back and forth between both domains, the first research question emerged: *"What are the aspects that drive and steer the design process in a collaborative environment?"* This research question (addressed in section 4) then led to the identification of research questions two, three, four and five. This is an example of the research approach described in this section, as the initial domain of research influenced the rest of the work. Research question two dives into design theory, question three into design process augmentation, question four into the creative capabilities of LLMs, and question five into how the conceptual descriptions can be translated into something tangible and practical. Finally question six reflects on the results of each question in the wider socio-economic context. As such, each question builds on the insights resulting from the previous question.

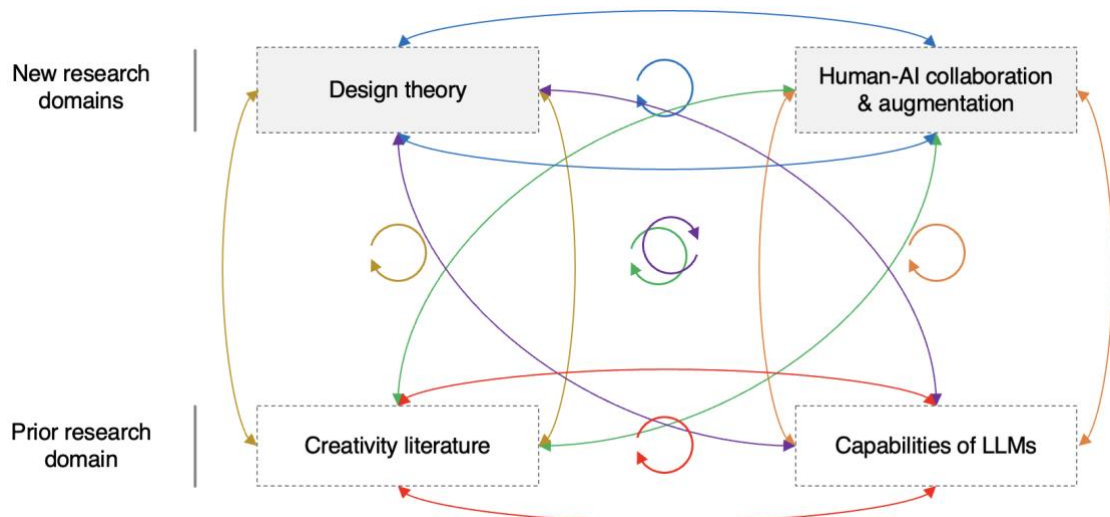


Figure 1: A visual representation of the literature review approach

### 2.3 Ensuring internal validity

To compensate for the potential lack of internal validity due to this thesis' iterative nature – something that inherently leads to propositions based on imperfect understanding and as such to potentially invalid premises – a focus group was held with a creativity theory and human-AI interaction expert. The objective of the focus group was twofold: 1) to validate the conclusions and findings, and 2) to explore the implications of the findings for the expert-related domains of research. The slide deck used for this focus-group is handed in as a separate deliverable. Given the fact that already existing conclusions had to be validated, the focus group was held in week fourteen of the total of 20 weeks of this research, such that enough relevant material could be presented. The insights derived from the focus group included 1) a set of questions surrounding the connections between various theoretical investigations and 2) pointed at the lack of a red thread that integrated all theoretical investigations into a single logical story that attends to the central research question of this thesis. These results have led to the restructuring of this thesis to its current form.

### 2.4 Ensuring practical relevance

Given the fact that the topic of this thesis is closely related to debates about human labor and the job market, it feels fitting to create a space for the interpretation of this thesis' results beyond their theoretical contributions. Therefore, research question six, addressed in section nine, reflects on the wider socio-economic implications of various ways in which augmentation systems can be deployed in the context of a capitalistic economy. Besides, this master thesis represents the end of the educational curriculum at the faculty of Industrial Design Engineering (IDE), and more specifically the Strategic Product Design (SPD) master program. As such, it seems adequate to reflect on one of the biggest (personal) challenges regarding the (strategic) design discipline: explaining to others what it is that a designer does, where design activities can be performed, and what the value of design is for organizations. Section twelve attends to this challenge by reflecting on the nature of (strategic) design, based on this thesis' theoretical contemplations, and its relation to different types of functions within various forms of organizations. As such, it translates theoretical findings to practical insights that may inform both novice designers as well as the management of the IDE faculty and SPD master program.

### 2.5 A note on the technological developments of LLMs

During the creation of this thesis, multiple new generations of LLMs have been introduced. For example, with regards to the frontier models, the successor to OpenAI's GPT-4 (OpenAI et al., 2024), GPT-4o, has been introduced. Even more recently, OpenAI introduced OpenAI o1<sup>7</sup>, an LLM based on a new training mechanism that enhanced its

---

<sup>7</sup> <https://openai.com/o1/>

‘reasoning’ capabilities. Although keeping up with the newest state of LLMs is important for the practical application of this thesis’ theory in today’s implementations, the broader objective of this research is to progress our understanding of the conceptual construct of augmented design with LLMs being a means to do so. As such, the definition of the types of design process augmentation has been mostly informed by the characteristics of the design process rather than the current state of LLM capabilities.

### 3. Defining LLM-based designer augmentation

---

This section aims to clarify what is understood as *design process augmentation* in this thesis. For this, Engelbart's (1962) framework for human intellect augmentation is used as it is the most unifying and rigorous thought framework for this particular issue (Xia & Maes, 2013).

---

To allow mankind to better face the increasingly complex world generated by its own manifestations, Engelbart (1962) argued that there was a need to conceive a conceptual understanding of how human intellect could be augmented or amplified. Paraphrasing from Engelbart (1962), human intellect augmentation is defined as increasing the capability of an individual to approach a complex problem situation, to gain comprehension, and to derive solutions to problems. The word 'better' implies 'speedier', 'of higher quality', and 'enabling what was not possible before' (Engelbart, 1962). This set of terms thus describes the metrics along which the level of output-based success of an augmented process are measured. The system in which humans operate and are augmented are described by four components, including *artifacts*, *language*, *methods*, and *learning* (Engelbart, 1962). This system is called the H-LAM/T system. Although this framework describes the intellect augmenting components, it doesn't yet provide an all-encompassing framework for studying an LLM-augmented human design process, as no dominating process or direction for the augmentation is provided.

A perspective that is able to translate Engelbart's framework for intellect augmentation to the topic of investigation of this thesis is the socio-technical systems (STS) theory belonging to the construct of sociomateriality (Fabri et al., 2023). The umbrella term of sociomateriality comprises the theory of the socio-technical system (STS) (Leonardi, 2013) that conceptualizes information systems as two interrelated subsystems: 1) the technical system that includes the *technology* and *tasks* required to produce products or services, and 2) the social system that describes the *people* and *structure* that together define the relationship between both systems (Döppner et al., 2019; Lyytinen & Newman, 2008). Instead of identifying the objects of intellect augmentation, this theory presents the components that constitute a hybrid human-computational system interacting via a certain *structure* to accomplish a *task* (Döppner et al., 2019). One could argue that what is defined as *structure* is, among things such as the relationship and interactions between the human and computational actors as described in (Döppner et al., 2019), a collection of *language*, and *methods* (from H-LAM/T), with the *technology* being a form of an *artifact* (Engelbart, 1962) (Table 1). In-

process *learning* then forms a characteristic of the human, and, depending on the technology, also one of the technology.

STS theory	H-LAM/T system
Human	Learning
Technology	Artifact (Learning)
Structure	Language, Methods
Task	-

Table 1: The relation between H-LAM/T and STS theory

Given the fact that the category of artifacts includes technology (Engelbart, 1962), our topic of interest, design process augmentation using LLMs, can be expressed in the H-LAM/T system. An LLM can be considered a form of an *artifact* that is used to augment the capacity of a designer given certain (existing) methods, via the modality of *language*. Expressed in STS theory, this research aims to conceptualize how LLMs, a *technology*, can augment a *human*, provided the *structure* in which the individual operates, to address a certain *task*. Whereas the H-LAM/T system provides a more detailed view on the types of intellect-augmenting points of leverage, the STS theory provides a more holistic account of the components that are part of the sociomaterial construct of human-technology hybrid processes.

LLM-based designer augmentation can thus be defined as “increasing the capacity of a designer to approach a complex problem situation, to gain comprehension, and to derive solutions to problems in a design project by introducing an LLM as a new type of artifact that complements the existing sociomaterial structure of design” (adapted from Engelbart, 1962). Before an account of designer augmentation can be established, it is necessary to conceive a framework of the design process in the context of human-AI collaboration that forms the foundation for further augmentation practices.

#### Takeaway box 1

LLM-based designer augmentation can be defined as “increasing the capacity of a designer to approach a complex problem situation, to gain comprehension, and to derive solutions to problems in a design project by introducing an LLM as a new type of artifact that complements the existing sociomaterial structure of design”



## 4. On the integration of design theory and human-AI collaboration frameworks

---

To design is to create new things that don't exist yet. Understanding how this process of conceiving the unknown can be augmented by LLMs requires both an overview of the factors that play a role in the collaborative process between a designer and LLM, and an account of the coverage of existing design theories relative to these factors. This leads to the first research question:

*What are the aspects that drive and steer the design process in a collaborative environment?*

---

### 4.1 Approach

First, literature about human-AI interaction and collaboration is introduced, as it establishes a wider scope on the factors that influence the actions of an actor, both human and computational, than individual design theories. This helps to identify a list of actor-related characteristics deemed critical for describing a process of collaboration between two or more actors, which is important for the creation of effective LLM-based design process augmentation. Second, several major design theories are discussed to identify their purpose and limitations, and to thus arrive at an initial view on each theory's coverage of the actor-related characteristics that define the actor's behavior in a collaborative setting.

### 4.2 Human-AI collaboration

To identify the characteristics and factors that affect the progression of a collaborative process between two or more actors, this subsection addresses various frameworks about human-AI collaboration and their limitations given the context of design. For example, the 'Symbiotic co-evolution in sociotechnical systems' framework (Döppner et al., 2019), and the Co-creative Framework for Interaction design (COFI) to design creative AI partners (Rezwana & Maher, 2023) aim to describe the dynamic interplay – or interactions – between humans and computational systems. Both frameworks however lack the detailed descriptive vocabulary for specifying a process of human-AI collaborative design. The former framework remains too high level and observative, describing only the spaces, such as the actor's actions, that evolve without further theoretical foundations to explain why certain changes occur within these spaces. The latter is more categorical in nature, synthesizing towards a selection interaction

patterns, derived from current human-AI collaborative systems, that can be used for interaction design. However, it doesn't provide enough theoretical grounding to explain the effect of these interaction patterns on the progression of the process.

Another framework is that from Fabri et al. (2023) who have created a taxonomy for human and AI-enabled hybrid systems. In line with the definition of design process augmentation from the previous section, this framework builds on the theory of sociomateriality (Orlikowski & Scott, 2008). The taxonomy makes a distinction between three types of entities: *human agent*, *AI-enabled system*, and *sociomaterial practices* (figure 2).

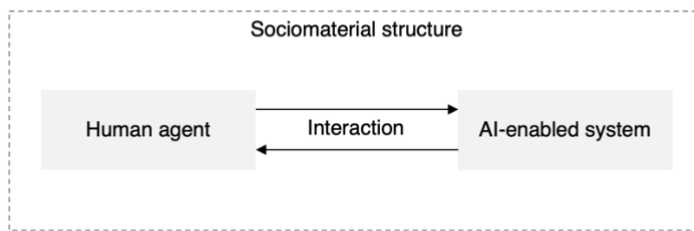


Figure 2: Taxonomy for human & AI-enabled hybrid systems

Table 2 highlights how the main categories of the taxonomy can be related to the STS theory (Leonardi, 2013) and the H-LAM/T system (Engelbart, 1962). The *human agent* can be positioned within the bucket of 'people', whereas the *AI-enabled system* corresponds to the descriptions of 'technology' and 'artifacts' of the STS theory and H-LAM/T system respectively. The *socio-material practices* comprise both the task and structure, including 'language', 'methods', and 'learning', as these entities define the interfaces that connect the human and AI-enabled system.

STS theory	H-LAM/T system	Taxonomy for human & AI-enabled hybrid systems
People	Learning	Human agent
Technology	Artifact (Learning)	AI-enabled system
Task		Sociomaterial practices
Structure	Language, Methods	

Table 2: Mapping the STS theory, H-LAM/T system, and the Taxonomy for human & AI-enabled hybrid systems

Within these three distinctive entities the taxonomy identifies 3 dimensions that describe the functions of the *human agent* and *AI-enabled system* (figure 3). Both human and AI entities are described by their 1) cognitive functions, 2) interaction characteristics related to the other entity, and 3) focus. The third entity, the *sociomaterial practices*, is described by the dimensions of 1) form of interworking, 2) mode of interworking, and 3) the type of learning that occurs. Each dimension is further

specified by various characteristics that can be present at various levels in each of the entities.

Layer 1: Sociomaterial entities	Layer 2: Dimensions		Layer 3: Characteristics									
Human (human agency)	Human cognitive functions	NE	Perceiving	Reasoning	Predicting	Planning	Decision-making	Explaining	Interacting	Creating	Empathizing	
	Interaction human to AI	ME	Facilitating			Verifying			Supplementing			
	Human focus	ME	Sensemaking		Creativity		Compassion		Flexibility			
AI (material agency)	AI cognitive functions	NE	Perceiving	Reasoning	Predicting	Planning	Decision-making	Interacting	Creating			
	Interaction AI to human	ME	Facilitating			Verifying			Supplementing			
	AI focus	ME	Automation				Augmentation					
Sociomaterial practices	Form of interworking	ME	Parallel			Sequential			Flexible			
	Mode of interworking	ME	Singular				Continuous					
	Learning	ME	None	AI learns		Human learns		Human and AI learn separately		Co-evolution		

ME = mutually exclusive, NE = non-exclusive

Figure 3: Taxonomy for human & AI-enabled hybrid systems (Fabri et al., 2023)

Although this taxonomy improves the clarity of the characteristics of potential synergies between humans and AI-enabled systems, it doesn't yet describe the (set of) processes(es) that accomplish(es) the task of the socio-technical system. For example, taking 'augmentation' (see figure 3, category: AI focus) as the focus of the AI entity doesn't yet describe the various ways in which the AI-enabled system can augment the human activity given a certain task – within this thesis the task is considered to be the design process. If the human focus would be on 'creativity' (see figure 3, category: human focus) – a relevant focus given the subject of this research being related to the process of design – the question arises how the AI could augment the human's creativity. Beyond agent functions, the question arises as to what mechanisms lead to the capabilities of both agents.

A framework that is more useful in describing these underlying mechanisms in the context of a collaborative process between a human and an AI-based system is the framework for human-AI collaborative design space exploration (DSE) (table 3), as it describes the characteristics of both human and computational agents that influence their action in a certain process (Viros-i-Martin & Selva, 2021). This framework originates from the discipline of systems engineering and describes the facets involved in a team of a human and a cognitive assistant (CA), where both agents are considered intelligent and able to learn, given their shared task space that allows for inter-agent interaction. Instead of starting with interaction qualities (as in COFI), this framework describes the characteristics of both human and AI agents, which then form the affordances of the interaction that is able to emerge in their shared design space.

Human-AI collaborative design space exploration framework			
Perception	Estimated state of the world	Human	An approximation of the state of the design task
		CA	[Not defined, implied to be similar to the human]
	Inferred state of the CA/human	Human	The expected interpreted state of the CA given the interaction
		CA	[Not defined, implied to be similar to the human]
Factors that constitute the state of the agents	Internal state (Dynamic during design process)	Human	Motivation, cognitive workload, situational awareness, <i>trust in CA</i> , <u>design preferences</u> , <u>the degree of design fixation</u>
		CA	Dialogue state, proactivity, role ('historian', analyst, explorer, expert, critic)
	Knowledge	Human	Domain, design process, problem specific
		CA	Problem database, heuristic database, historical designs
	Goals	Human	Design, learning, collaboration
		CA	Design, learning, collaboration (alleviating design fixation, increasing common ground and trust in CA)
	Attributes (Static during design task)	Human	Cognitive style, level of expertise, risk aversion, expectations, preconceptions
		CA	Learning, intelligence, autonomy, communication
Legend			
- <i>Italic</i> = Related to human machine collaboration literature*			
- <u>Underlined</u> = Related to design literature*			
*(Viros-i-Martin & Selva, 2021)			

Table 3: A framework for human-AI collaborative design space exploration (Viros-i-Martin & Selva, 2021)

Paraphrasing the content in table 3, the framework describes how, initiated by an actor's perception that comprises an estimated state of the world and inferred state of the CA or Human (depending on the perspective), mediated by its knowledge, driven by its internal state, steered by its goals, and constrained by its attributes, two actors can both take action. This is also visualized in figure 4. This framework thus introduces an overview of contextual aspects relevant for the collaboration between two human and computational actors.

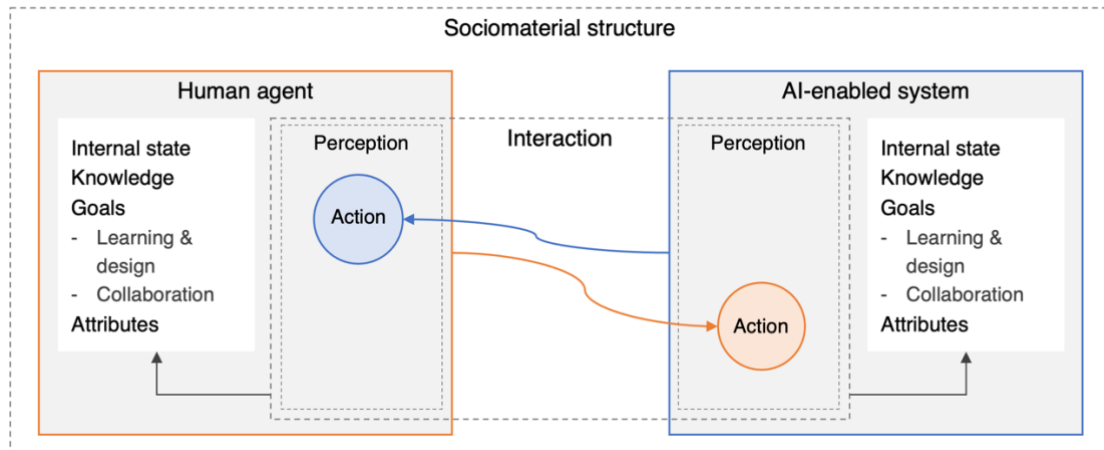


Figure 4: A visual representation of interaction between two actors

Although these and many other studies (e.g. Döppner et al., 2019; Kim & Maher, 2023; Lin et al., 2023; Liu et al., 2024; Rezwana & Maher, 2023; Schwartz et al., 2023; Sharma et al., 2024; Tsiakas & Murray-Rust, 2024) have aimed to construct a framework that captures the interaction of human and computational system characteristics, including various components of human and artificial or cognitive agents (CAs) and their shared design space, there is a limited understanding of the application of such theory in the domain of design theory. More specifically, these frameworks aim to describe the categories and levels of interaction and collaboration between human and computational agents but lack the detail to describe and explain how a collaborative pair of designer and AI progress through the design process. Given the interest in establishing a more detailed view on the design process that incorporates the contextual aspects identified by Viros-i-Martin and Selva (2021), the question arises how the designer's internal state, knowledge and goals (learning, design, collaboration) affect the design process? The next subsection provides an initial view on the relation between existing design theories and these actor-related characteristics.

### Takeaway box 2

There are many frameworks for describing the construct of human-AI interaction. The most fitting framework given the objective of this thesis is the collaborative human-AI design space exploration framework (Viros-i-Martin & Selva, 2021). It presents a set of actor-related characteristics that both affect and are affected by the actions of the actor itself or the other actor in the collaborative process. These characteristics form the basis for understanding behavior in collaboration. As such, the following section aims to find the overlap between existing design theories and this list of characteristics that drives the design process.

## 4.3 Design theory

If one were to visit design theory literature to search for a perspective on the design process that could be used as a blueprint for understanding how aspects such as a

designer's perception, goals, knowledge, and internal state (Viros-i-Martin & Selva, 2021) would affect the progression of the design process, one would face a range of design theories, each from a different viewpoint and with a different limitations. Below, the four main theories of design<sup>8</sup> are briefly discussed to highlight the lack of integration and identify their relation relative to the actor-related factors from the collaborative human-AI design space exploration framework (Viros-i-Martin & Selva, 2021). Moreover, table 4 shows to which actor-related characteristics each individual theory of design is connected in its descriptions and explanations of the design process.

First, the Concept-Knowledge (C-K) theory of design (Hatchuel & Weil, 2002) provides a view on the relation between one's knowledge and the concepts that are able to emerge from it. However, it lacks a time-based framework of knowledge progression through design. Second, another theory, problem-solution (P-S) co-evolution in design (Maher & Poon, 1996), does provide such a time-based framework of design progression driven by epistemic perceived uncertainty (Cash et al., 2023), but is not conceptually grounded in an account of knowledge representations (Crilly, 2021a, 2021b). Third, the (situated) FBS framework (Gero, 1990; Gero & Kannengiesser, 2004), presents a view on design process progression, the types of content that allow a designer to progress, and the way in which new insights can emerge through the expectation externalization and by interpretation based on one's space of memory. However, there is no connection between the space of memory and the space of knowledge described in the C-K design theory. Fourth, the Uncertainty-Driven-Action (UDA) model of design establishes a connected account of various layers that constitute a designer's behavior, connecting a designer's uncertainty perception, motivation, goals, and actions that result from it (Cash & Kreye, 2017). However, just like the situated FBS framework and P-S co-evolution model of design, it lacks the integration of an account of the designer's knowledge. Four important observations can be made with regards to this discussion of design theories and table 4

- First, there is no integration between the design theories. This is also indicated by the lack of between-theory references<sup>9</sup>.
- Second each theory only covers a part of the actor-related characteristics as defined in the human-AI design space exploration framework (Viros-i-Martin & Selva, 2021) (table 3). The C-K theory only covers the 'knowledge' category. P-S co-evolution focusses on the internal state of the designer (perceived uncertainty, while only loosely referring to 'knowledge' (the light-grey fill color in table 4 highlights this loose connection). The situated FBS framework is mostly descriptive and disconnected from any actor-related characteristics. However, it does provide a description of the effect of an actor's 'perception' in

---

<sup>8</sup> These four theories were discussed during the focus group without further suggestions for other design process related theories.

<sup>9</sup> An observation shared during the focus group.

the *interpreted world* space on the progression of the design process. Although not explicitly described in the theory itself, this can also be used as mechanism for inferring the state of the CA, hence the use of the light grey color. Finally, the UDA model covers the internal state (uncertainty perception) and goals.

- Third, design theory describes the design progression from the perspective of the human cognition but is not represented in the computational domain. As such, there is a limited perspective on the ability of LLMs to augment the design process.
- Fourth, the category ‘attributes’ is not accounted for in each of the design theories. Although these attributes are important in shaping an effective collaborative process, they can be regarded as too specific to and dependent on the individual relative to this thesis’ objective to further discuss them.

Human-AI collaborative design space exploration framework			C-K theory	P-S co-evolution	(Situating) FBS framework	UDA model
Perception	Estimated state of the world	Human				
		AI				
	Inferred state of the CA/human	Human				
		AI				
Factors that constitute the state of the agents	Internal state (Dynamic during design process)	Human				
		AI				
	Knowledge	Human				
		AI				
	Goals	Human				
		AI				
	Attributes (Static during design task)	Human				
		AI				

Table 4: The coverage of design theories relative to the framework from Viros-i-Martin and Selva (2021)

#### 4.4 Towards a unified design process framework for augmentation

To address the first and second observation, one could argue that the most fundamental factor that currently hinders both the creation and the interpretation of current design process augmentation systems, is the lack of conceptual connection between design theories, as these theoretical divisions result in a partial and thus incomplete connection between descriptions of the design process and descriptions of the factors that are said to affect the behavior of an individual in a collaborative process. This observation is in line with the proposals for conceptual expansion of design theory beyond the individual (Crilly, 2021b). Either the disconnection between theories, or the lack of conceptual expansion of a single theory leads to a gap in conceptual understanding of the factors

that influence the progression of a human designer through the design process in a collaborative manner with an LLM-based augmentation system. To address the gap of partial mappings between individual design theories and the human-AI collaborative design space exploration framework, it is necessary to unify the design theories mentioned in this section to create a single time-based design process framework that covers the aspects as mentioned in table 4. This leads to the second research question:

*How should the design process be conceived such that it allows for the identification of the types of augmentation, and the points of interaction that manifest these augmentation practices?*

The next section (section 5) elaborates on each of the discussed design theories to work towards a unified design process framework. The resulting framework then forms the basis for identifying not only how the design process can be augmented on a content-level, but also the ways in which the interaction that leads to the augmentation can be shaped. Besides, this framework can also serve as a conceptual map for explaining and understanding existing and future empirical findings. After all, the unification of the above-mentioned design theories connects a designer's cognitive activities, grounded in knowledge, to his externalized expectations, or propositions, embodied in actions. Here, actions can be regarded as both the points where the actors interact, as well as indicators of the actor's internal state. This leads to the third research question:

*How can the design process be augmented?*

Section 6 addresses this question while building on the unified framework to identify how this process can be augmented according to the definition of human intellect augmentation from Engelbart (1962) (section 3)

#### **4.5 Towards the integration of AI-based systems in design**

To address the third observation about the representation of design theory in the computational domain, an account of the design capabilities of GenAI-based systems, such as Large Language Models (LLMs), should be constructed. This leads to the fourth research question:

*What are the capabilities of LLMs to realize design process augmentation practices?*

Section seven provides a theoretical analysis of the ability of LLMs to augment the design process, and, as will become clearer in the section itself, an emerging view on the capabilities and (current) limitations of LLMs to simulate a process of design themselves. Building on this perspective on LLM capabilities to realize design process augmentation, section eight introduces an interface based on the conceptual framework of design process augmentation. This section attends to the fifth research question:



*How can design process augmentation practices be manifested in practice?*

This interface not only represents the findings of this thesis in a format that is easily interpretable but also acts as a boundary object for the technological capabilities of LLMs in design. The objective of this boundary object is to stimulate the discussion and reflection on what *can* be possible and what *should* be possible; hence, it forms the basis for reflection on the ethical side of the use of AI in the design practice. This reflection is executed according to the sixth and final research question of this thesis:

*What are the socio-economic implications of LLM-based augmentation systems?*

### **Takeaway box 3**

Currently, the landscape of design theory is hindering the exploration of both the creation and interpretation of human-AI collaborative design process, as multiple theories exist that each describe a different perspective on the design process. The effect is that one can only partially map between individual design process theory-related cognitive functions and the actor-related characteristics described in human-AI collaboration literature. This hinders the ability to understand how the design process can be augmented by AI-based augmentation systems. This thesis therefore aims to 1) unify the design process theories, 2) define how the design process can be augmented, 3) what the capabilities of LLMs are to realize such augmentation practices, and 4) to design an interface that realizes the resulting conceptual framework of LLM-based design process augmentation. These four research domains correspond to the second, third, fourth and fifth research question respectively. The sixth research question aims to shine a light on the socio-economic implications of various forms of augmentation systems.

## 5. A unified design process framework

---

As a designer, it is a fascinating experience to observe how through time and effort one can create new things based on existing things. This seemingly magical process of letting a new artifact emerge from existing matter, where the artifact may show some resemblance to the things it is based on but also introduces many new things, is the subject of investigation in design and creativity literature. Design research is specifically interested in understanding how designers design, and why they act the way they do (Cash et al., 2023). Various design theories have been conceived to address these questions of *how* and *why*.

The previous section highlighted four theories and showed how each theory has its own focus and limitations. Given the aim of this thesis to augment the design process, and as such to understand how this construct can be manifested in a collaborative process between two or more actors, it was argued in the previous section that the existing design theories have to be unified to enable the integration of design theory and literature on human-AI collaboration. This section therefore elaborates on each of the discussed design theories to work towards this unified design process framework. It thus attends to the second research question:

*How should the design process be conceived such that it allows for the identification of the types of augmentation, and the points of interaction that manifest these augmentation practices?*

---

### 5.1 Approach

The following sections elaborate on the previously listed design theories. The approach to the theoretical contemplations is informed by two motivations that can be derived from the research question stated above. First, the framework has to describe the process from such a perspective that it becomes clear how the design process can be amplified or augmented. To attend to this motivation, this section's approach builds on the C-K theory that argues that concepts provide an unknown proposition to one's knowledge, and that one's knowledge is thus the key informant for both creation and interpretation in design (Hatchuel & Weil, 2002). As such, the working hypothesis for the remainder of this section is that design process augmentation fundamentally starts

with amplifying the knowledge of a designer. The design process theories will thus be united by grounding them in a common representation of knowledge.

Second, the framework should depict a time-based process that shows the points where the augmentation system can interact with the design process of a designer. The approach, that concentrates on this second motivation, is to use the model of P-S co-evolution as a basis for unification, since it describes a time-based model of design progression and has the most overlap with other theories (see the horizontal overlap in table 4). Especially the latter argument makes it the most suitable fundament for unification, as it connects the Uncertainty Driven Action Model (UDA) of design (Cash & Kreye, 2017) that describes a designer's actions, and thus interaction points with the external world, to the C-K theory that forms the theoretical foundation for the working hypothesis of how to augment the design process.

As can be read further down, the common ground of knowledge representations is found in 'schemata' theory that provides a theoretical fundament for representing knowledge in memory. This last notion of 'memory' makes it especially suitable to the integration of the situated FBS framework in the unified set of design theories, since this theory has grounded an individual's interpretations and expectations in the space of memory (Gero & Kannengiesser, 2014). Schemata theory is thus able to link the situated FBS framework to the design theories grounded in knowledge.

The process of constructing the unified design process framework can be described in two main parts:

*Part 1: Grounding design theories in knowledge (section 5.2 – 5.10)*

The first part attends to the first motivation underlying this section's approach. This part unifies the three main design theories while connecting them to knowledge representations. The result is a framework that is capable of expressing (cognitive) problem-solution co-evolution (Cash et al., 2023; Maher & Poon, 1996) in terms of both the situated FBS framework (Gero & Kannengiesser, 2014) and Concept-Knowledge theory (Hatchuel & Weil, 2002). Here, the time-based model of design as P-S co-evolution is taken as a starting point for design theory unification based on a common knowledge representation.

*Part 2: Identifying interaction points for augmentation (section 5.11 – 5.16)*

The second part attends to the second motivation underlying this section's approach. This part adds the layers of behavior from the Uncertainty-Driven Action (UDA) model of design (Cash & Kreye, 2017). This leads to the formulation of a three-stage mental iteration which highlights the three points of interaction between the designer and an external augmentation system.

## *Part 1: Grounding design theories in knowledge representations*

### **5.2 Design as co-evolution of problems and solutions**

One of the most noteworthy views on the design process is that design is an iterative process in which problem and solution spaces co-evolve, where the designer explores both the conceptual spaces of problems and solutions (Wiltschnig et al., 2013). Problems are situations that are unfavorable and require change towards a desired state (Crilly, 2021b). Solutions are the vehicles that could realize that change. This view was introduced by Dorst and Cross in 2001 (Dorst & Cross, 2001), who drew on Mary Lou Maher's computational implementation of the biologically inspired co-evolution model for describing how problems and solutions in design evolve individually while also having a reciprocal effect on one another (e.g. Maher, 1994; Maher & Poon, 1995; Maher & Poon, 1996). In design co-evolution model presented by Maher & Poon, there are four different types of problem-solution transitions; from problem to solution (P-S); from solution to problem (S-P); from solution to solution (S-S); from problem to problem (P-P) (Maher & Poon, 1996). A progression from P-S to S-P implies that in the former transition the solution is viewed as a solution to the problem, whereas in the latter that same solution introduces or proposes a new problem.

### **5.3 Cognitive co-evolution model**

In their cognitive co-evolutionary model of design, Cash et al. (2023) have not only described how the process of design proceeds but explained through a cognitive lens what initiates and manages this process. They describe how meta-cognitive monitoring and control, relative to a designer's knowledge and understanding of the world, leads to co-evolutionary transitions (e.g. from problem to solution (P-S), S-P, S-S, and P-P) and progression (e.g. from P-S to S-P), via a concept that is often called epistemic or perceived uncertainty (Cash & Kreye, 2017; Christensen & Ball, 2018) (Cash et al., 2023). This epistemic or perceived uncertainty is previously defined as a *"perceived lack of knowledge by an individual, in the form of deficiencies in any stage or activity of the process that can be characterized as not definite, not known, or not reliable"* (Cash et al., 2023, p.5). In doing so, they define the 'direction' of the co-evolutionary process in design as: *"direction of cognition and behavior based on a designer's perceived uncertainty and associated metacognition."* (p. 5). Given this state of meta-cognitive perceived uncertainty, the individual aims to reduce the uncertainty by exploring gaps by proposing candidate ideas.

Although Cash et al. (2023) do not explicitly mention knowledge as a space that co-evolves, they implicitly argue that it is a central component in the design process, as the state of it relative to problem or task at hand defines the level of perceived uncertainty that drives the design process. Knowledge can thus be seen as a component that is more fundamental than problems and solutions as it guides the individual's interpretation of ideas as either being problems or solutions. If we were to explain how

designers can be augmented by LLMs, it is thus critical to move beyond the notion of design as problem-solution co-evolution. The next section will further detail this observation by discussing the previous notions of design as problem-solution co-evolution.

#### 5.4 The origin & evolution of theories about problem-solution co-evolution

In the original interpretation of co-evolution, inspired by the biological process expressed in genetics, the genotype of an animal represented the information representation that formed the basis for generating design solution candidates, whereas the phenotype (an animal's observable traits) was considered a proposition based on the state of knowledge represented by the genotype (Crilly, 2021b; Maher & Poon, 1996). The analogy extended the models of computational design creativity from that time, as not only the *design propositions* in the *design space* were considered to evolve due to mutations and other actions, but also the *fitness-function* within the *performance space* that selected the idea propositions. Here the changes in both the *design propositions* and *fitness function* are a direct result of changes in the underlying information representations, or genotypes, resulting from the proposition in the design space. Both design and performance spaces were later renamed into problem and solution spaces by Maher and Poon (Crilly, 2021b); the space of design propositions (phenotypes) was equated to solutions, and the performance space with its fitness-function to the problems.

Another interpretation is that of John Langrish who uses the terms 'selectemes' and 'recipemes' (2004). Following the formulation of self-replicating ideas that are subject to processes of mutation, combination and selection as 'memes' (Dawkins, 1976), he describes 'selectemes' as ideas about problems, requirements and needs, and 'recipemes' as ideas about possible ways to solve the problems and satisfy the requirements (Langrish, 2004). The selectemes and recipemes introduce a more conversational nature of the interaction between problems and solutions, and somewhat resemble the notion of the creative process as 'blind variation' (ideas for solutions) and 'selective retention' (ideas of problems and requirements) (Campbell, 1962). In contrast to Campbell's model of creative thought, however, both the recipemes and selectemes co-evolve. For example, as new recipemes introduce new dimensions that are currently not accounted for in the selectemes, the selectemes adapt to the newly embodied dimensions in the recipemes, such that these dimensions steer the next iteration of ideas about solutions. In short, this interpretation of problem-solution co-evolution focuses more on the co-evolution of the mechanisms or processes that lead to the identification of the problem and solutions (i.e., the recipes (recipemes) and selection (selectemes) that allows the actor to progress its creative thoughts). Although beyond the scope of this thesis, an interesting comparison can be made between the mechanism of recipemes and selectemes and the way in which Generative Adversarial Networks (GANs) work.

Two observations can be made. First, it is important to note that in the process of relating various frameworks, the genotypes, the set of information representations, are lost in translation as they are not part of the definition of either the problem or solution space. This is especially critical considering transitions between ideas as problems and ideas as solutions. Revisiting the definition of a problem, one could argue that a problem isn't simply an unfavorable state of your environment or design solution (Crilly, 2021b), but an interpretation of it given certain presupposed expectations. The question then becomes, what initiates the transition from an idea at  $t=i$  as a solution – hence resulting in a high fitness value at  $t=i$  – to that same idea at  $t=i+1$  as a problem – presenting a low fitness value at  $t=i+1$ ? The interpretation of an idea as either problem or solution can be attributed to an individual's understanding of its environment or design solution relative to its knowledge, where understanding is defined as one's ability to identify the application and limits of one's knowledge (Cash et al., 2023). Similar to how knowledge and understanding influence one's perception, one's knowledge also informs the solutions that are able to emerge. Hatchuel and Weil mention that *“the space of concepts C cannot be dissociated from the space of knowledge K as its definition is relative to K”* (Hatchuel & Weil, 2002, p. 11 & 12). One could argue that the model of co-evolution in design is thus not complete without an account of one's knowledge state (Crilly, 2021b) – the information representations (genotypes).

Second, equating the space of problems to the 'performance space' (Crilly, 2021b) is problematic as both constructs represent different things. The performance space is an account of the perceived fit between an idea (or solution, phenotype and recipeme) and a fitness-function that aims towards a certain objective. Inherent to the definition of a 'problem' is the low fitness value of an idea; it is a state in which the idea can reside that is relative to the objective and not an absolute characteristic inherent to the idea. Similarly, Crilly argues that instead of looking at problems and solution occupying different spaces and thus being of a different kind, we have to regard both a problem and a solution as the same idea that resides in two different states or is perceived through two different lenses (Crilly, 2021a). Problems and solutions are two different sides of the same coin. Instead of representing different spaces for problems and solutions, Crilly introduces a heterarchy of ideas (figure 5), where ideas both govern and are governed by each other, hence representing a problem or a solution respectively depending on the perspective one takes (2021).

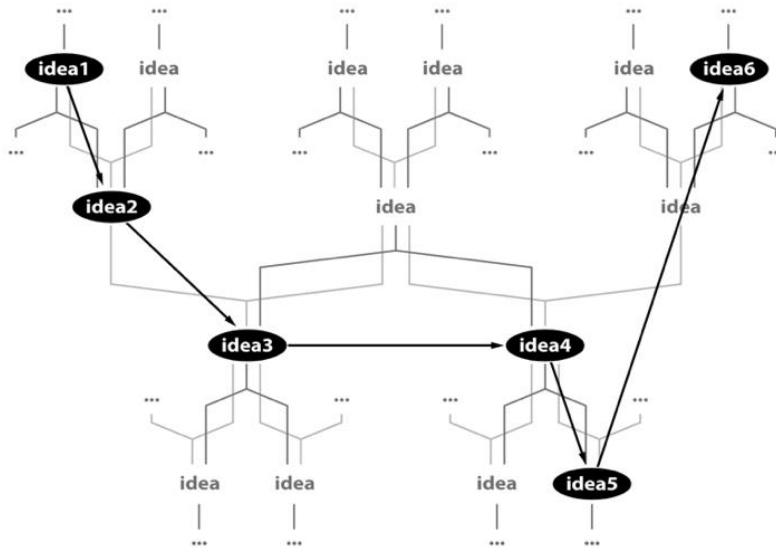


Figure 5: Co-evolution visualized in a heterarchy of design ideas (Crilly, 2021a)

These observations point at some theoretically inelegant characteristics of the problem-solution co-evolution model of design. Although the model, that describes the process of resolving the apparent tension between problems and solutions spaces through their mutual adaptation, sounds intuitive at first, it presents some challenges when trying to explain the apparent transitions between the two. One could argue that instead of co-evolving ideas embodied as either problems or solutions, a distinction has to be made between the evolution of one's ability to identify both the problem and solution in the presented idea from the ability to come up with the actual idea itself. As such, it is critical to construct better definitions of both *ideas*, as well the *evolution* of one's view on an idea. Following the logic of Cash et al. (2023) and Crilly and Moroşanu Firth (2019) while relating back to the original conception of biological co-evolution expressed in genotypes and phenotypes (Maher & Poon, 1996), the evolution of an idea should be grounded in an individual's knowledge, as knowledge is the most fundamental component that drives both processes of perception and creation (Hatchuel & Weil, 2002, p. 11 & 12). Building on the distinction between the idea and one's perspective on that idea, while aiming to contribute to a more elegant notion of co-evolution in design, the next two sections describe the different perspectives on ideas and their interaction with the set of 'known entities' or knowledge and their relations that define one's view on those ideas.

### 5.5 Idea representations: Function, Structure & Behavior (FBS)

To arrive at a more nuanced perspective on ideas, another influential concept within design literature will be introduced; that of the FBS ontology that describes the various functions of design prototypes as Functions, Behaviors and Structures (Gero, 1990).

- Functions (F) are defined as the set of premises a design solution should satisfy, and as the relation between the goal of a human and the behavior of the system (Gero, 1990; Takeda et al., 1990).
- Structures (S) are then introduced as an intermediate step or proposition toward a design solution, describing at a design component level the composition of the components that realize the set of functions. Takeda et al. call these structures ‘design solution candidates’ consisting of a set of attributes, that represent the properties of the design solution candidate (1990).
- Finally, the concept of ‘behavior’ is two-fold; behavior is conceived as both the ‘expected behavior’ ( $B_e$ ), describing the “*the syntax by which the semantics represented by function can be achieved*” (Gero, 1990, p. 28), and the ‘structure’s behavior’ ( $B_s$ ) which refers to the actual structure’s functioning.

Following the example given by Gero (1990), functions of a window can include the provision of daylight and control over ventilation. Structures are the glazing, frame, and their respective properties.  $B_e$  is the internal representation of the materialization of the described set of functions, such as the amount of daylight and ventilation provided and thus the set of attribute levels that are assumed in the functions.  $B_s$ , or actual behavior, is the actual amount of daylight and ventilation provided by the structure (or design solution candidate (Takeda et al., 1990)).

#### **5.5.1 Transitions between idea representations**

In the FBS model, eight different processes are described that together form the progression of a designer through various design iterations, as visualized in figure 6 (Gero & Kannengiesser, 2004). A process can be either transformational or comparative. Only process four, *evaluation*, is such a comparative process.

1. *Formulation* is the transformation of an initial set of functions into expected behavior  $B_e$  that is required to enable this function
2. *Synthesis* is the process of transforming the expected behavior  $B_e$  into a structure that is capable of satisfying  $B_e$
3. *Analysis* captures the structure’s actual behavior  $B_s$
4. *Evaluation* compares the expected behavior  $B_e$  and actual structure behavior  $B_s$
5. *Documentation* presents the description of the design given the current state of the functions, structure and behavior
6. *Reformulation of structure (S) variables* changes, resulting from a discrepancy between  $B_e$  and  $B_s$ , changes the composition of the structure
7. *Reformulation of the expected behavior  $B_e$* , resulting from a discrepancy between  $B_e$  and  $B_s$ , changes the composition of the structure
8. *Reformulation of the set of functions (F)*, resulting from a discrepancy between  $B_e$  and  $B_s$ , changes the set of functions



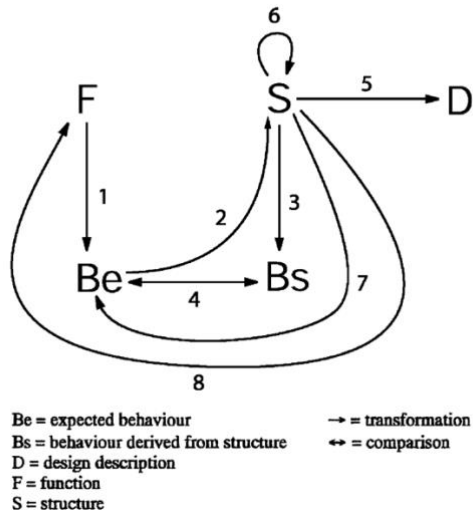


Figure 6: The FBS framework (Gero & Kannengiesser, 2004)

To find  $B_s$ , an individual needs to externalize this representation into an external environment through what Cash and Kreye (2017) call ‘representation action’; action undertaken by an individual to improve its understanding of the design space through representational exploration.  $B_s$  is then the result of the confrontation of that *structure* with a set of dimensions, or *affordances* (Glăveanu, 2013), from an external environment. Although the FBS ontology makes an implicit distinction between an internal and an external world, it doesn’t yet exhaustively address the conceptual nuances that this distinction demands.

### 5.5.2 Situated FBS framework

In the ‘situated FBS’ framework (Gero & Kannengiesser, 2004), the concept of a dynamic environment in which the FBS framework is ‘situated’ is introduced. Besides reformulating ‘internal’ into ‘interpreted’ world, a third type of ‘situatedness’ is identified and defined as the ‘expected world’ that is positioned within the interpreted world. It is an instance of the interpreted world obtained through what is described as ‘focusing’ (Gero & Kannengiesser, 2004). They argue that a change in one of the three worlds may bring about changes in the other worlds and vice versa (figure 7) (Gero & Kannengiesser, 2004).

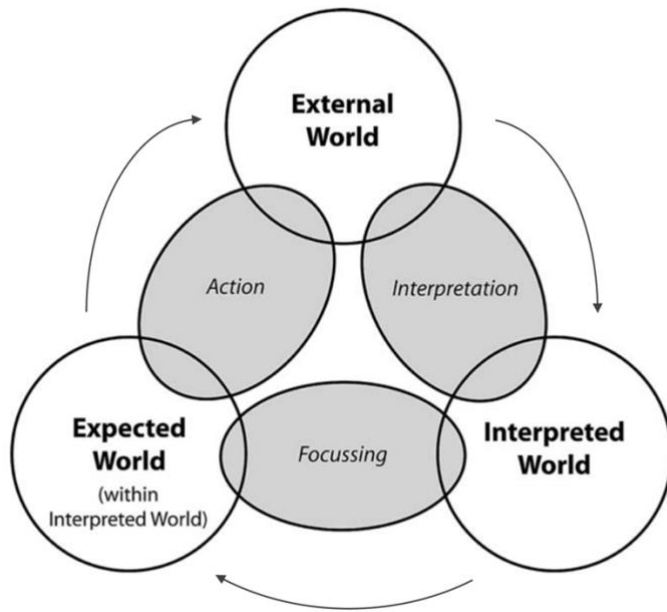


Figure 7: Situatedness, adapted from Gero & Kannengiesser (2004)

To understand how changes in either of the three worlds emerge and affect the others, two fundamental concepts that underly the situated FBS need to be understood: *situatedness* and *constructive memory* (Gero & Kannengiesser, 2004). *Situatedness* is concerned with the inclusion of a designer's environment within the process of design. By acting in the world, a designer changes its environment. Those changes are then perceived by the designer and influencing the course of its following actions. There is a recursive process between doing and perceiving (Schon & Wiggins, 1992) where *interpretation* of the interpreted and external world (figure 7) plays a central role. *Constructive memory* concerns the notion of dynamic memory, co-constructed during the action-interpretation cycles by both newly incoming perceptions and previous states of memory (Gero & Kannengiesser, 2004). *Constructive memory* operates both at the intersection of the external and interpreted world and within the interpreted world following a 'push-pull process'. First, the external world provides new data to the interpreted world (push), while the interpreted world perceives the new data through a previous experience-based lens (pull). This is what Gero & Kannengieser call the 'original experience' (2004). Second, the original experience participates in a push-pull process with the actor's internal memory. Besides introducing a distinction between internal (interpreted and expected) and external worlds, the situated FBS framework therefore implicitly adds another space that is assumed to co-evolve: the space of memory. Here the experience represents the data in the form of an interpreted representation of an external representation that 'pushes' new content to the memory, whereas the memory 'pulls' specific parts of that 'experience' based on previous accounts of memory.

Based on these two concepts a 'situated FBS framework' (figure 8) is introduced that situates instances of Function, Structure and Behavior into the three identified worlds. The eight processes from the original FBS framework are expanded to a total of twenty processes to describe the additional transitions between worlds. Besides the introduction of additional processes, two other observations can be made. First, 'R', representing the original design requirements from the design process' offset, is introduced as the artifact that initiates the process of interpretation between external and internal world. Second, the 'focusing' process type is introduced, complementing the processes of push-pull, transformation and comparison. This process enables the designer to formulate a design state space, a subset of the *interpreted* or conceptual (Hatchuel & Weil, 2002) space, within which the design solution is searched (Gero & Kannengiesser, 2004). Within this subset, or *expected* world, assumptions are made about possible functions, structures and behaviors that satisfy R. Below, the relation between the processes from the original FBS framework (figure 7) and the situated FBS framework (figure 8) is explained.

- Process 1, *formulation*, (figure 7) is partitioned into process 1 to 10 (figure 8); the direct mapping from function to expected behavior is divided in processes that bridge the different worlds (external, interpreted and expected) through interpretation, constructive memory and focus.
- Process 2, *synthesis*, is divided into processes 11 and 12, where expected behavior ( $B_e^i$ ) is first transformed into an expected structure ( $S_e^i$ ) before it is transformed into an externalized version of that structure ( $S^e$ ).
- Process 3, *analysis*, comprises processes 13 and 14. Here, the externalized structure ( $S^e$ ) is first interpreted through the push-pull process between the external and internal world and within the latter resulting in an updated version of ( $S_e^i$ ), before being transformed into an updated version of the set of interpreted behaviors ( $B_e^i$ ).
- Process 4, *evaluation*, remains the same as the original framework already made a distinction between actual and expected behavior. It is visualized by process 15.
- Process 5, *documentation*, created by translating S into D in the original framework, has been redefined as externalizing the state of  $F_e^i$ ,  $B_e^i$ , and  $S_e^i$  to  $F_e$ ,  $B_e$  and  $S_e$ .
- Process 6, *reformulation type 1*, comprises process 6, 9, and 13. The situated FBS framework introduces more nuance to the ways in which descriptions of *structure*, or more specifically  $S_e^i$ , can be revised by the designer. First, the designer may interpret the external world differently (process 13). Second, the process of constructive memory can identify different aspects in  $S_i$ , as such altering the interpretation of structure. Third, the instances of *structure* included in the *expected world* may be changed.
- Process 7, *reformulation type 2*, is described by process 5, 8, 14, and 19. It follows the same logic as *reformulation type 1* (process 5, 8, and 19), where process 14

depicts the transformation of the interpreted structure ( $S_i$ ) to the interpreted behavior ( $B_i$ ).

- Process 8, *reformulation type 3*, is described by processes 4, 7, 16, 20. It follows the same logic as *reformulation type 1* and 2 (process 4, 7, and 20), where process 16 depicts the transformation of the interpreted structure ( $B_i$ ) to the interpreted behavior ( $F_i$ ).

As previously mentioned, the model of *situatedness* makes a distinction between the *external* and the *interpreted world*, where the latter is described as a space of concepts defined as variables transformed through interpretation (Hatchuel & Weil, 2002). Within the *interpreted world*, the model of *constructive memory* introduces the notions of both *experiences* and *memory*. Although the situated FBS framework describes the various perspectives on ideas, satisfying the first request for better a better conception of ideas, it doesn't yet describe from a bottom-up perspective the knowledge structure and composition in which the interpretations based on *constructive memory* are embedded. As such, it cannot yet exhaustively describe the transition between ideas as problems and solutions and the well-known concept of *emergence*, defined as the process of uncovering something existing (finding a solution by expressing it to the external world) as well as being presented with something new (interpreting the external representation and finding new dimensions) (Dorst, 2019), as there is no account of an individual's knowledge on which the interpreted world is based. The next section introduces the Concept-Knowledge design theory (Hatchuel & Weil, 2002) that relates the space of knowledge to the situated FBS framework and thus emergence.

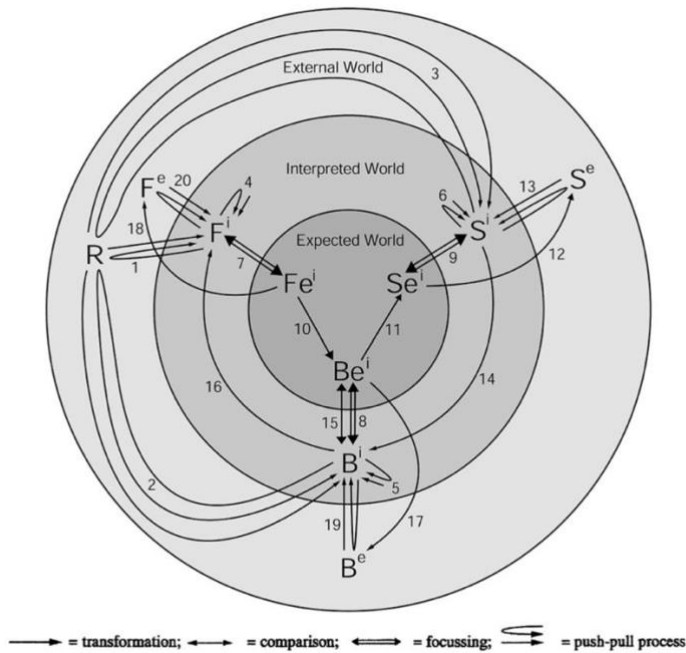


Figure 8: Situated FBS framework (Gero & Kannengiesser, 2004)

## 5.6 Ideas, concepts, and knowledge: Concept-Knowledge design theory

To describe how transitions between perceptions of ideas as problems or solutions occur, it is critical to understand how the knowledge structure from which ideas originate relates to the various perspectives on ideas offered by the situated FBS framework. Whereas the situated FBS framework does not explain the fundamental structure from which these ideas arise, the C-K design theory provides an emerging perspective on the relation between design manifestations and knowledge. Similar to the description of *emergence*, Hatchuel and Weil (2002) argue in their Concept-Knowledge (C-K) design theory that “*a concept evokes an unknown proposition relative to the knowledge available*” (p. 11). They argue that design reasoning must therefore always make a distinction between the spaces of concepts and knowledge, as a concept can only be partially understood given the state of knowledge at the time of expressing the concept (Hatchuel & Weil, 2002). This observation also resonates with the concept of meta-cognitive perceived uncertainty described in (Cash et al., 2023), describing how an individual is driven to explore the design space by a perceived lack of knowledge about the space of the design project. Hatchuel and Weil (2002) define knowledge as a set of logical propositions with a logical status while concepts are conceived as propositions without logical status, and thus as elements that are not part of one’s knowledge (Hatchuel & Weil, 2002). Although one could challenge this view by building on the theory of epistemic logic probabilities (van Eijck & Schwarzentruher, n.d.), this is beyond the scope of this thesis and will as such not be covered. Here, we assume the observation of Hatchuel and Weil on the topic of logic relative to knowledge and memory (2022) to be correct as it will not challenge the remaining content in a major way.

Although the C-K theory presents a perspective on knowledge, it does not yet provide a connection to *constructive memory* from the situated FBS framework. In other words, there is no unified conception of the fundamental structures underlying these terms. Moreover, the situated FBS framework doesn’t yet provide a structure or representation that visualizes the progression of memory contents. As such, no clear perspective on design as a composite construct of knowledge, memory, concepts, and the interaction between these spaces exists that can describe how transitions between ideas in problem or solution states occur. Therefore, both theories have to be reconceived based on the same fundamental building blocks. An additional theoretical benefit of understanding the construct of knowledge relative to C-K design theory and the FBS ontology is the direct link to the model of cognitive co-evolution that describes how a perceived lack of knowledge results in the perceived uncertainty (Cash et al., 2023). To relate the C-K design theory to the situated FBS framework it is necessary to understand how the spaces of knowledge and concepts relate to the external, interpreted and expected worlds as defined in the situated FBS framework. To do so, schemata, knowledge representations, will be used as building blocks that can express both theories.

### 5.7 Schemata as knowledge representations

Just like the FBS framework introduces various representations of ideas, as being functions, structures or behaviors, Rumelhart and Ortony (2017) introduce *schemata* (singular: schema) as representations of knowledge. Schemata are “*data structures for representing the generic concepts stored in memory*” (Rumelhart & Ortony, 2017, p. 101). They form the structure abstracted from and afterwards underlying the memory components perceived through perception. A *schema* is built up out of variables and subschemata, the latter including its own variables and subschemata. Figure 9 visualizes a simplified version of a conceptual schema using the Active Structural Network (ASN) representation (Norman & Rumelhart, 1975). Variables can take on different values that are constraint to both inform the individual on the type of objects that can be related to the schema variables, as well as to function as a foundation for the prediction of the most suitable variable values if not enough information is presented. Variable levels have to be regarded as distributions of possible values. As such, in the case of too little information, an individual might prefer to choose a variable level that is closer to the average of the distribution. On the contrary, when the environment presents a situation that is difficult to capture in the given set of schemata, more deviant values can be chosen to interpret that situation (Rumelhart & Ortony, 2017). The process of comprehension is then described as having the perceptions activate a set of schemata that are then used to understand the scenario by choosing appropriate variable values, such that the schemata accurately reflect the situation. Schemata can also describe an individual’s lack of understanding, the case in which no schema is readily available for the interpretation of a specific scenario.

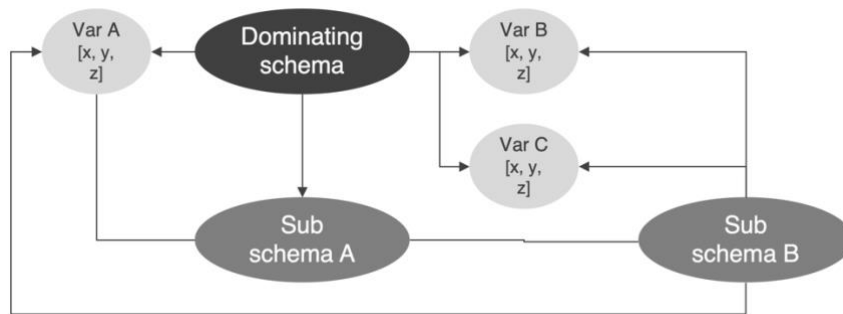


Figure 9: An example of a schema (simplified from Rumelhart & Ortony, 2017)

### 5.8 Instantiated schemata, Situated FBS framework & C-K theory

To relate the situated FBS framework to the C-K theory, schemata theory has to be able to express both *concepts*, artifacts in the interpreted and expected worlds, and *knowledge*. Interpretations – another way to describe concepts in the interpreted world – are considered the offshoot of the comprehension process, also called *instantiated schemata* (Rumelhart & Ortony, 2017). *Instantiated schemata* are schemata where the variable values are chosen and thus fixed. What is regarded as the interpreted world in the situated FBS framework (Gero & Kannengiesser, 2004) and the space of concepts in

the C-K theory (Hatchuel & Weil, 2002), is thus the result of *comprehension* expressed in terms of *instantiated schemata*.

How then does the *original experience* that is part of the *constructive memory* as described in (Gero & Kannengiesser, 2004) relate to this? As visualized in figure 10, The original experience (E) is defined as the result of incoming new data pushing and prior experiences pulling data in. One could argue that this is a process similar to that of one's senses perceiving and thus pushing in new data that starts the process of associating relevant schemata, while the selected schemata also inform the identification of other relevant subschemata. Experience creation is thus the process of selecting the appropriate set of schemata based on both associations informed by incoming new data as well as the schemata structures themselves representing the starting point for identifying relevant subschemata. Memories (M) are then constructed by prior schemata instantiations predicting (pulling) the variable values from the newly selected schemata based on contextual similarity (Rumelhart & Ortony, 2017), as well as variable values being selected based on their ability to explain the present context (figure 10). These observations are in line with research on *experience* and *memory*, arguing that experiences are the essence of memory (Tulving, 1989). In other words, the fundament for memories is the set of 'known' schemata (knowledge) that are identified as related to the perceived context (experience). Building on the observation that memories are interpretations (Rumelhart & Ortony, 2017), it can be thus stated that the interpreted world of the situated FBS framework (Gero & Kannengiesser, 2004) is filled with memories, defined as instances of knowledge.

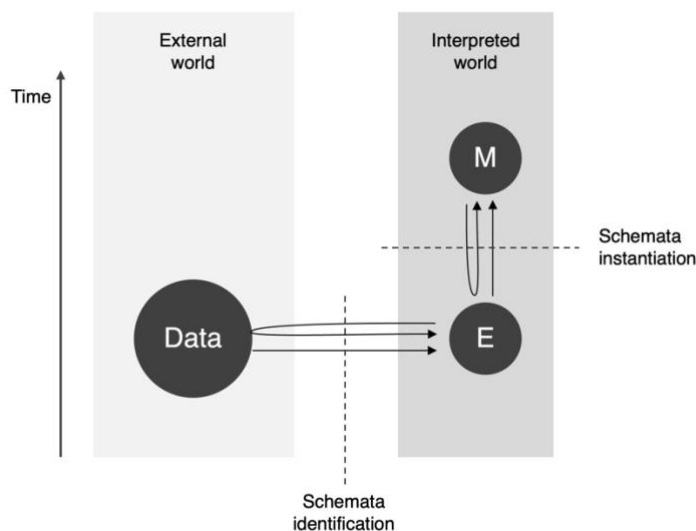


Figure 10: Constructive memory and schemata (adapted from Gero & Kannengiesser (2004))

Where then can the space of concepts be positioned? Memories, being instantiated schemata with fixed variable values chosen from their respective distributions, should

according to Hatchuel and Weil (2002) be considered a form of concepts, as the space of concepts is defined as variables transformed through interpretation resulting in propositions without logical status. Combining this with the observation that the interpreted world is the space of memory, one could argue that memory and concepts are two terms that describe the same thing: instantiated schemata (figure 10). Following the definition of Gero and Kannengiesser, the expected world is a subset of concepts or memories, where assumptions are made about possible functions, behaviors and structures that satisfy 'R' (2004). In short, it can be described as the 'design state space'. There is no structural difference between instantiated schemata in the interpreted and expected world, besides their perceived level of capacity to positively contribute to the task at hand. For example, using the example from Gero (1990) of a window design, one could argue that instantiated schemata related to a 'road bike', are not perceived as being relevant considering the task of designing a window. As such, the 'bike-schemata' will not be selected as candidate for the expected world. To further refine our conception of the expected world, let's consider instances of *function*, *structure*, and *behavior* expressed in terms of *schemata*. Again, the example used in (Gero, 1990) of a window design is used.

- An example of a window's *function* ( $F_e^i$ ), defined as a premise the design should be able to satisfy (Gero, 1990), is the provision of daylight. This function is captured in an active structural network representation as visualized in figure 11 and can be considered the dominating schema. This also aligns with the order of the situated FBS processes within the expected world that start with the definition of the *function*.
- The *structure* ( $S_e^i$ ), the composition of the design solution at a component level (Gero, 1990), is a set of subschemata within the dominating schema that contribute to the overall *function*. In this case, the subschema 'transfer' further details the structural elements and their variables (figure 11).
- The *behavior* ( $B_e^i$ ), defined as "*the syntax by which the semantics represented by the function can be [or are expected to be] achieved*" (Gero, 1990, p. 28), is the expected value of variable A. This expected variable value is the result of assumptions about variables B – E. These variable values are either chosen based on the variable levels from prior instantiated schemata similar to the schemata being associated with the current design task, or chosen such that they represent the average of the variable distribution in the case of limited related memories (Rumelhart & Ortony, 2017).

The behavior ( $B^i$ ), which is the interpretation of an externalized *structure* ( $S^e$ ) can potentially introduce new schemata that were previously not associated with the design task. An example could be the interior color of the room, a variable within the subschema 'inside', in which the window will be positioned. When that room is painted black, the expected behavior resulting from a selection of certain window dimensions might not correspond with the actual perceived light levels as the *variable* 'color' was not yet taken into account.



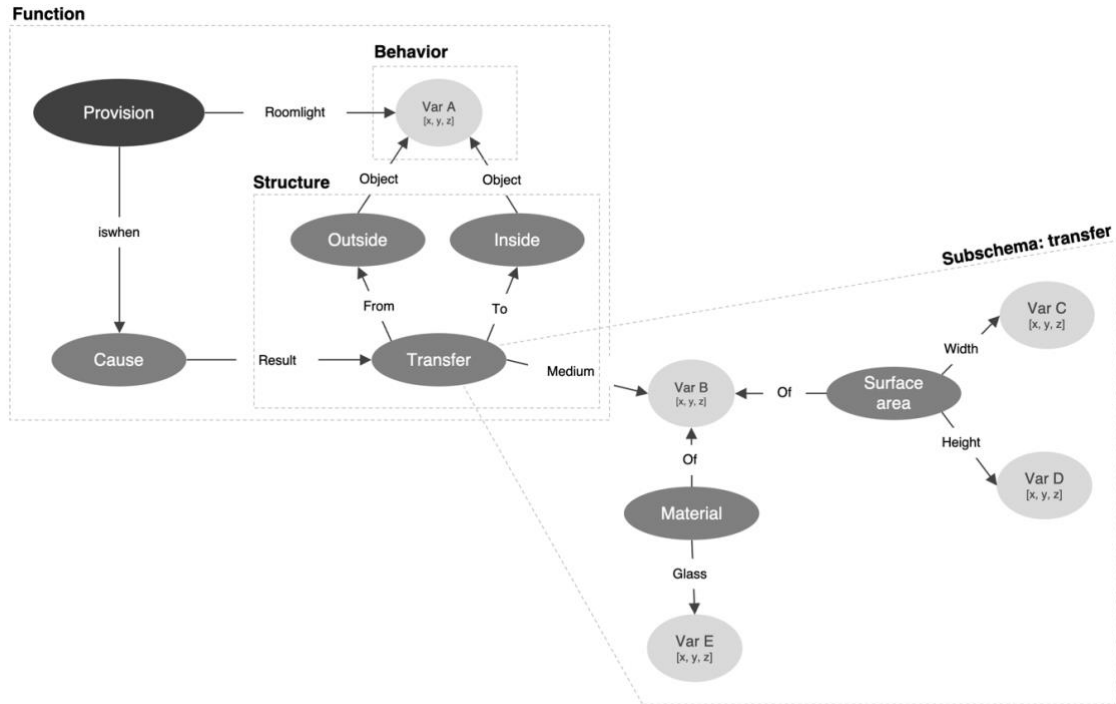


Figure 11: Active structural network representation of window function

### 5.9 ‘Problems’ and ‘solutions’ expressed in a unified design process framework

The observations on conceptual integration of the situated FBS framework and C-K design theory, based on schemata as knowledge representations are visualized in figure 12. Recalling our initial call for better definitions of both *ideas* and *transitions* between ideas grounded in a common representation of *knowledge*, it can be said that ideas, either as problems or solutions, and the transitions between them can now be expressed in the situated FBS framework that is connected to the C-K design theory (Hatchuel & Weil, 2002) via schemata. The next section connects the model of design as problems-solution co-evolution by expressing that model in terms of the alteration of schemata as a result of the processes defined in the situated FBS framework.

As previously stated, the model of the design process as the co-evolution of problems and solutions presented trouble in describing what underlying phenomena are changing during co-evolutionary transitions between problems and solutions. As such, it was argued that a division had to be made between the idea itself, and the ability of an individual to identify either the problem or solution in it. The situated FBS framework (Gero & Kannengiesser, 2004) then describes the various perspectives on ideas in design (*functions*, *behaviors*, and *structures*) relative to their ‘world space’, where the C-K design theory (Hatchuel & Weil, 2002) in combination with the *schemata* as knowledge representations (Rumelhart & Ortony, 2017) provides more detail to the knowledge structure that forms the basis for these ideas. Building on the theory proposed by the

situated FBS framework, both problems and solutions can be expressed in the components of the other two frameworks (figure 12). Here, the process of *evaluation* between *expected interpreted behavior* ( $B_e^i$ ) and *interpreted behavior* ( $B^i$ ) can be seen as the key to understanding the different perspectives an individual can have on the same idea based on a certain state of knowledge (figure 12).

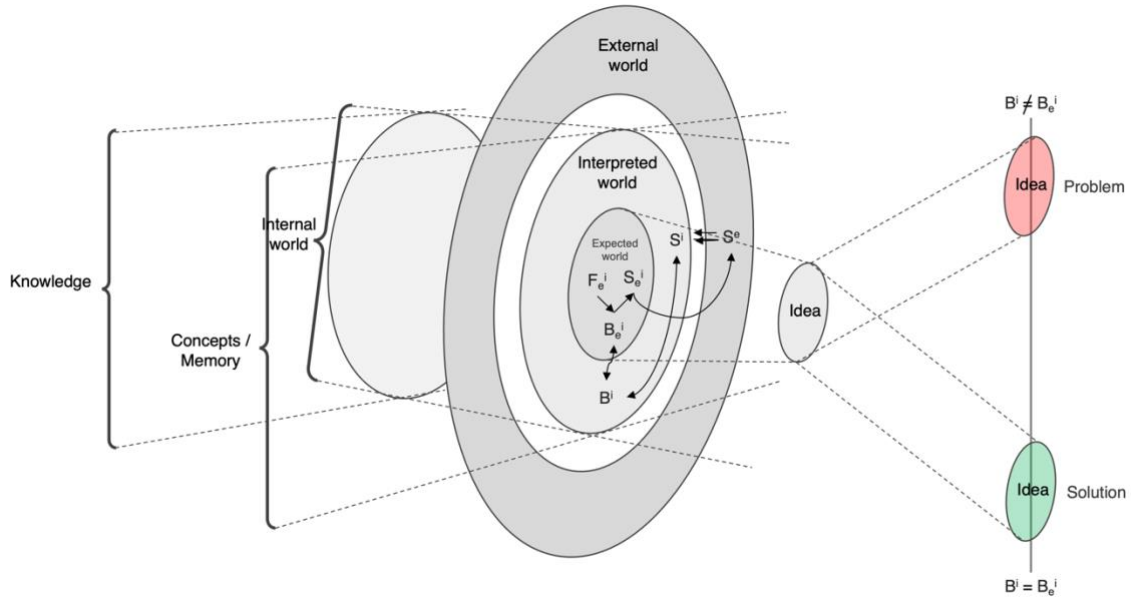


Figure 12: Spaces that co-evolve in design and their relation to problems and solutions

To understand how this works, let's consider a time-based framework of design progression as visualized in figure 13. With regards to the spaces, there are two co-evolving spaces on the highest level of abstraction: the *external* and *internal* world. Within the internal world three 'nested' spaces (Crilly, 2021b) are identified along two different axes. First, the *concept/memory* and *knowledge* spaces (Hatchuel & Weil, 2002) can be identified along the axes that divide the spaces with logical status from those without. Second, within the space of *concepts/memory*, the *interpreted* world can be distinguished from the *expected* world (Gero & Kannengiesser, 2004) to make a division between the complete set of associated instantiated schemata triggered during interpretation and those identified as critical to better understand during a next design iteration. Visualized below the external and internal world, the spaces of *problems* and *solutions* are depicted as derivate states in which one's thoughts about a specific idea can reside.

For each transition between P-S, S-P, P-P, and S-S, a zigzag across the external and internal world spaces can be observed. The artifact within the external world, 'R' (the initial design brief, or any externalized iteration of the design process) is interpreted within the space of concept/memory, mediated by the space of knowledge.

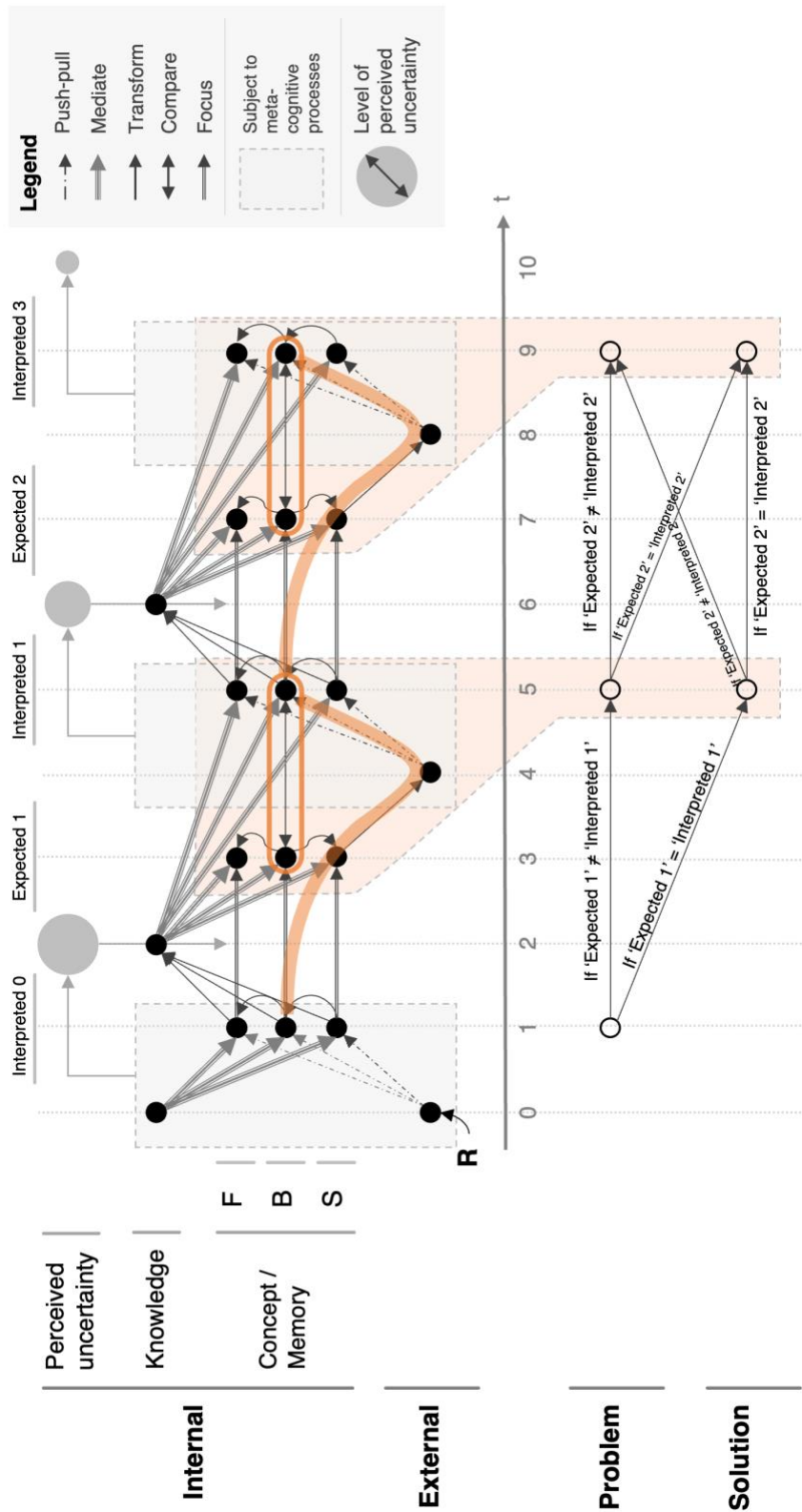


Figure 13: Transitions of problem/solution relative to external/concept/ knowledge spaces

Knowledge is regarded as a mediating variable as the process of comprehension – getting to an interpretation – is said to comprise a combination of perception and schemata activation. Since schemata represent knowledge, knowledge can be considered a mediating component in realizing the interpretation. During this zigzag, the concept space acts as the space of negotiation between an individual's knowledge and the affordances of its social and material environment, hence the external world. The process of comparing  $B_e^i$  and  $B_i$  (the rounded orange boxes) presents that negotiation and determines the direction of the transition from either a problem to a solution or to another problem. When  $B_i$  resembles  $B_e^i$ , the externalized expectation is interpreted as a solution. When  $B_i$  does not resemble  $B_e^i$ , it is perceived as a problem.

#### Takeaway box 4

To consolidate the findings from the previous sections, and to paraphrase figure 13, to design is to expand one's knowledge and understanding through proposition (from expected to external world), interpretation (from external to interpreted world), and evaluation (comparing between interpreted and expected world). Existing knowledge and memory are used to interpret and propose (*interpret* and *expect* in situated FBS terms), while being updated through evaluation afterwards. The quality of a design concept, whether imaginative or physical is thus as good as the understanding of the designer.

### 5.10 The relation between and progression of understanding and knowledge

To answer the question how an individual progresses between interpretations and projected expectations in the form of externalized ideas, the cognitive co-evolution model highlights the importance of meta-cognitive perceived uncertainty that results from an individual's lack of understanding (Cash et al., 2023). This leads to the question how *understanding* can be understood relative to one's knowledge?

According to Grimm (2014), one's understanding is one's ability to construct propositions of object and effect. He defines *understanding* as 'knowledge of causes'. It differs from knowledge as it is propositional; it is a chain of factual components ordered based on causal relations to describe how object A might result in effect B (figure 14). Here, 'object A' may be conceptualized as 'R' (from the situated FBS model) or any next version of an individual's externalized conception of the design brief. According to Grimm, building these causal chains of components can happen at various levels of detail. One's knowledge is the critical factor in the level of detail of such chains. For example, a child might understand that faulty wires can lead to a house being burned down, but a fire chief may understand every detail and step in between those two things leading to a more detailed and accurate representation of the event (Grimm, 2014).

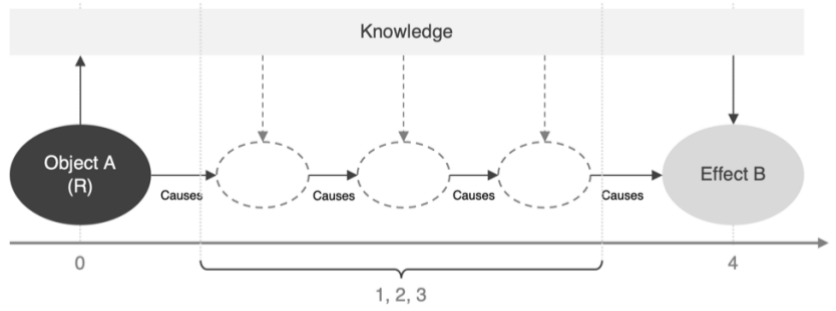


Figure 14: Understanding as knowledge of causes

How is this relevant to the discussion about design progression? The answer is that the *structure* from the FBS ontology can be understood as being one's understanding given a certain desired behavior, where the identification of the limits of one's understanding are said to lead to perceived uncertainty (Cash et al., 2023). Not knowing something results in the conclusion that one doesn't have a certain knowledge component. Not understanding something, means that either a) a knowledge component is missing, b) there is a lack of causal relation between knowledge components, or c) both. As such, *understanding* is the process of relating knowledge components such that they establish a certain effect or outcome; expressed in terms of FBS components, it is the *structure* by which a certain *behavior* is realized.

An illustrative active structural network representation is given in figure 15 to show how knowledge of causes can be understood relative to schemata. The object of interpretation, effect B, represents a *behavior* for which a *structure* of causally related knowledge components must be constructed that is able to describe it. Whenever a designer is not capable of generating this *structure* (e.g., when there is a lack of understanding), it results in a sense of perceived uncertainty. As such, perceived uncertainty can be conceived as a state resulting from the inability to construct a causally correct and sufficiently detailed set of knowledge components able to explain the *behavior* of the object that is being interpreted. Put differently, an idea as a *problem* and an individual in the state of *uncertainty* are both derivatives of a lack of understanding defined as knowledge of causes.

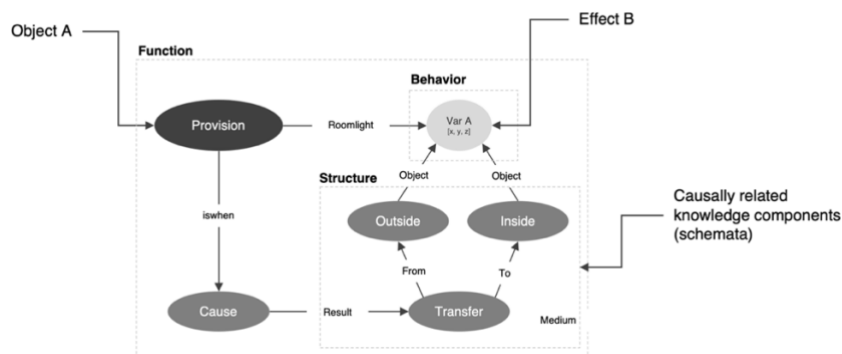


Figure 15: Understanding visualized as an Active Structural Network (ASN)

Based on a certain understanding or ‘knowledge of causes’, a new proposition is created within the interpreted world that is then externalized into the external world. In the case of the model presented in figure 13 the progression from the interpreted to expected world within the space of concepts, described as *focusing* by Gero & Kannengiesser (2022), can be regarded as an individual generating partial propositions of *structure* to fill in the gap between object and effect (figure 16), or function and behavior.

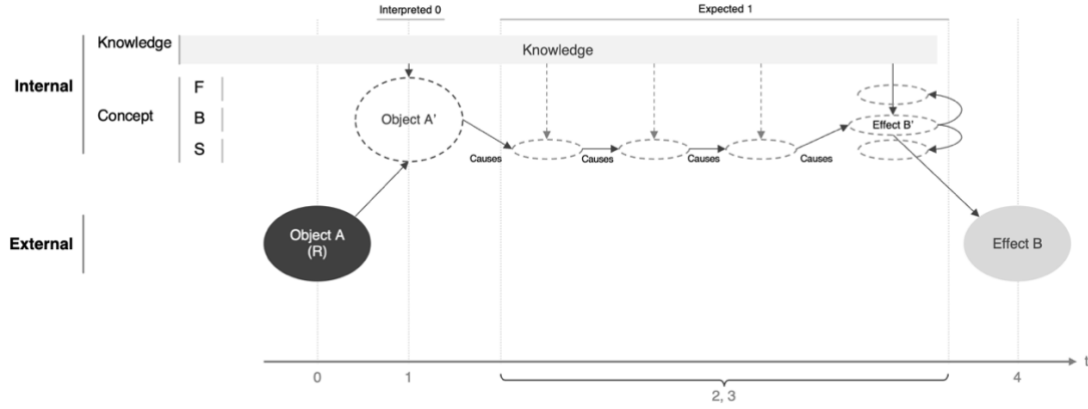


Figure 16: Knowledge of causes

The validation of this proposition based on incomplete knowledge occurs by interpreting the idea (and its effect), based on the knowledge from  $t=2$  and the conceptual memory from  $t=3$ . Due to the choice for schemata and variable values being dependent on imperfect knowledge and memories,  $B_i$  can be dissimilar to  $B_e^i$ , leading to a) another instance of perceived uncertainty, b) a slight transformation of the knowledge structure (by changing the schemata composition and/or changing the variable value distributions), and c) the observation that the idea is a problem. When there is no difference between  $B_i$  and  $B_e^i$  it can be assumed that the relations between knowledge components, or schemata, are both causal and complete. In other words, the distribution of variable values and the set of prior instantiated schemata were the right material for constructing a correct representation of an idea, leading to the idea being considered a solution.

Recalling earlier theories of co-evolution (Maher & Poon, 1996), the fitness function that resides in the performance space, later called the ‘problem space’, implies that the transition from problem to solution can be gradual and doesn’t have to be as black and white as is often suggested by the time-based visualizations of problem-solution co-evolution as in (Maher & Poon, 1996). Based on  $B_e^i$  and  $B_i$ , a gradual account of ideas as either problems or solutions (figure 13) can be established by formulating a function that captures the apparent fit between expected and interpreted behavior:

$$F(B_i, B_e^i) = -\|B_i - B_e^i\|$$

Given the fact that  $B_e^i$  and  $B_i$  can a) be expressed in terms of schemata with variable value distributions and b) in causally related subschemata composed in a context-dependent *structure*, the evaluation of the ‘fitness’ of the expected instantiation of the idea in the interpreted world relative to the idea as interpreted afterwards has the potential to be formalized. That however is beyond the scope of this thesis. For now, it is relevant observe that co-evolution of problems and solutions (P-S) can be regarded as a derivative process of the co-evolution of knowledge and concept/memory, with meta-cognitive perceived uncertainty being the result of a) the limited availability of necessary knowledge or b) limited established causal relationships between available knowledge components.

## *Part 2: Identifying interaction points for augmentation*

### **5.11 Towards interaction points between designer and augmentation system**

So far, a description of the design process has been conceived that describes design progression from the perspective of the spaces that are affected and driven by epistemic perceived uncertainty. Here, epistemic uncertainty is influenced by the identification of the limits of one’s understanding based on one’s knowledge. At a high level, the design process framework as conceived so far describes the spaces of concepts and knowledge.

Referring to the central research question of this section, the design process framework as visualized in figure 13 should not only describe the facets that can be augmented (by augmenting an individual’s knowledge), but also describe the points where interaction can occur between the designer’s internal cognition and another actor – for example, an LLM-based augmentation system – that manifests such augmentation. The remainder of this second part builds on the working hypothesis that the point where both actors are able to perceive and thus interact with each other is where an actor externalizes the outcome of its cognitive processes into an action that is then received by the other actor(s). The question then becomes what the relation is between the design process framework as conceived so far and the actions that result from it. The Uncertainty-Driven-Action (UDA) model of design (Cash & Kreye, 2017) presents a way forward.

#### **5.11.1 Layers of behavior**

The Uncertainty-Driven-Action (UDA) model of design activity proposed by Cash and Kreye (2017) describes the different layers that together constitute a designer’s behavior. Cash and Kreye (2017) argue that design activity is composed of three main levels: the *activity*, *task*, and *action* levels (figure 17). First, the *activity* level comprises motivation-driven and as such goal-led meta-tasks that initiate further behavior. An

example *activity* is designing a novel concept (Cash & Kreye, 2017). Cash and Kreye state that multiple *activities* can be held simultaneously by a designer, as such opening the possibility for the definition of design, learning and collaboration activities, following the categorization of *goals* of the framework from Viros-i-Martin and Selva (2021). Second, the *task* level includes a sequence of more specific processes that relate to the activities. Third, the *action* level describes the set of generic actions that map to processes from the task level. These actions can be described in terms of their cognitive processes (Cash & Kreye, 2017), and include three types: *representation*, *information* and *knowledge sharing* action.

Let's consider the design brief of designing a window again. Given this example, the following could be said about design activity at the three distinct layers:

- Activity level:
  - The design goal to deliver the best work possible for your client may lead to the definition of the overall motivation-led task to, for example, optimize the esthetics of the window in the environment in which it is placed.
  - The learning goal, which is aligned with the design goal (Viros-i-Martin & Selva, 2021) – this is also supported by the mental iteration design process framework, as it describes design as a process of proposition-based learning – is to understand the design conditions as well as possible
  - The collaboration goal is to provide contextual information around the actions taken by both actors, by ensuring alignment on goals (activity layer), and the task that is currently being executed (task layer)
- Task level: A task or process that contributes to this goal-led task is to find the optimal dimensions of the window relative to the wall in terms of esthetic balance.
- Action level: An action that contributes to finding these optimal dimensions could be to measure the wall in which the window is placed.

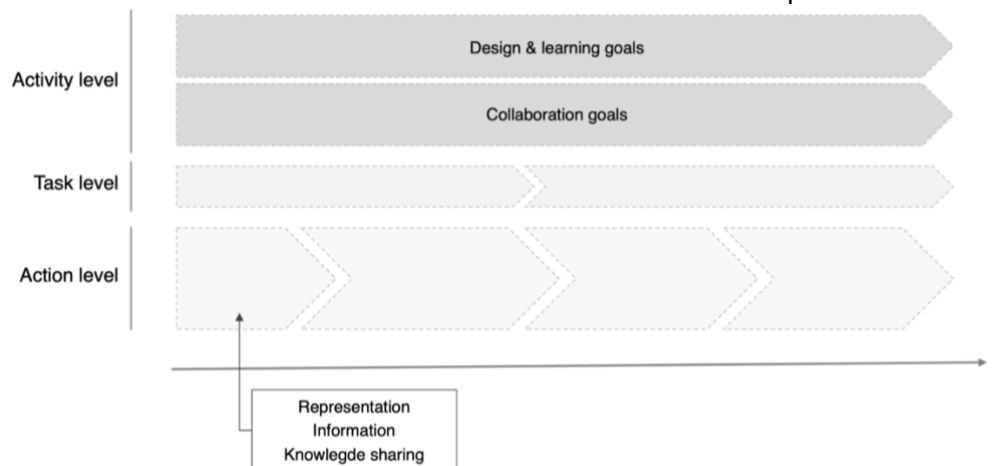


Figure 17: Layers of design activity



### 5.11.2 Types of design actions

To better understand how the three types of design actions relate to the three types of goal-led activities, let us now consider the action level in more detail. Cash and Kreye (2017) have identified three types of design actions positioned in the *action* level: *representation*, *information* and *knowledge sharing* action.

1. *Representation* action can be described as dealing with the perception and manipulation of external representations of information (Cash & Kreye, 2017, p. 21), where representations can also be regarded as simulations.
2. *Information* action is defined as “*dealing with data parts and their manipulation*” (Cash & Kreye, 2017, p. 19; Court, 1997). It focusses on finding data parts relevant for the interpretation of the design project and formation of the right conceptual understanding, and is driven by a perceived need, which is linked to a perceived uncertainty perception (Cash & Kreye, 2017). Furthermore, it is associated with cognitive processes describing the identification, manipulation, and transformation of data parts (Cash & Kreye, 2017).
3. *Knowledge sharing* action refers to the act of communicating to other team members driven by the objective to arrive at the same collective understanding of a certain problem (Cash & Kreye, 2017). It is highly dependent on effective communication between individuals, as knowledge sharing is part of an interpersonal exchange.

Let us now consider the relation between these types of actions and the design process framework as conceived in (see figure 18). First, *Information* action is about working with and manipulating data and can thus be described as acting at the intersection between the spaces of knowledge and concepts, where data parts are “*sought, reasoned about, and stored*” (Cash & Kreye, 2017, p. 19). This resembles the process that is used to interpret an external artifact and to, based on that interpretation, make a proposition (manipulation). Expressed in terms of schemata, *information* action can be defined as searching for schemata and proposing schemata structures to understand or interpret an artifact. To manipulate it, the logic of and behind a set of instantiated schemata is projected onto another set of instantiated schemata, as to propose a causally related composition of schemata that has a chance to be a correct representation of reality. This describes in terms of schemata what others call design patterns, analogies, or Boden’s creativity types/paradigm relatedness as detailed in figure 26 from section 6.5.

Second, *Representation* action both precedes and succeeds *information* action, as it can be equated to transition from the external to the internal world within the design process framework. First, it precedes *information* action as perception informs the formulation of a conceptual representation of the external artifact. Second, in succeeding *information* action, *representation* action is about externalizing an internal thought process, something that can be done via multiple media including, among others, language, drawings, or images. Here, *representation* action is the projection of

the expected composition of instantiated schemata onto a certain medium for the purpose of communication and further interpretation.

Third, *Knowledge sharing* action refers to one's explanation about the reasoning that led to either a specific interpretation of, or expectation about the data parts. As Cash and Kreye (2017, p. 20) put it: "*This type of action is associated with cognitive processes describing how knowledge is expressed with respect to an individual's understanding and beliefs*" (Court, 1997). Expressed in terms of schemata, *knowledge sharing* action is the descriptive expression of the logic or 'why' behind the composition of the schemata composition.

Since *representation* action bridges internal and external worlds, the implication is that in collaboration, both the initiation of an actor's perceived uncertainty through interpretation, as well as the externalization of *information* and *knowledge sharing* action emerge through *representation* action. In practice this means that every design action of a designer is either a combination of 1) *information* and *representation* action, 2) *knowledge* and *representation* action, or 3) a combination of all three actions. For example, if a designer chooses to go with a certain set of window dimensions, this choice resembles *information* action as it is a proposition created by manipulating the data parts from one's knowledge. As soon as he externalizes this choice, the designer participates in an act of *representation*. A designer's reasoning behind these choices can be regarded as *knowledge sharing* action. Again, when this reasoning is externalized, it is represented through some form of medium. Knowledge sharing action thus resembles the expression or description of one's understanding, also described as one's knowledge of causes (Grimm, 2014).

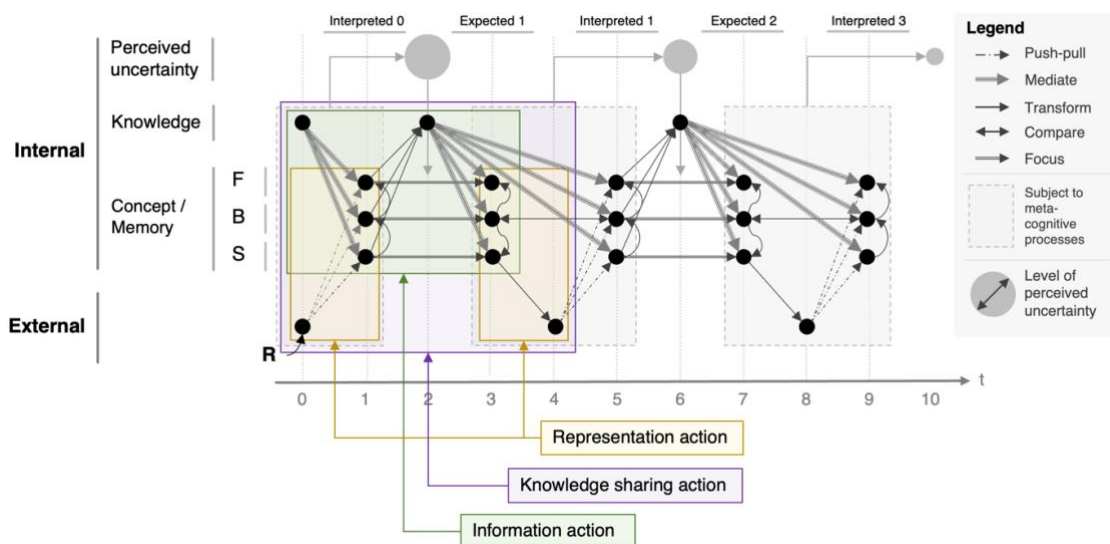


Figure 18: The relation between types of design actions and the design process framework

### 5.12 The relation between types of goals and design actions

The distinction between these types of design actions is relevant as they serve different purposes and contribute to different goals. With regards to the purpose of design actions, *information* action can be seen as mostly serving the individual that aims to alleviate a certain uncertainty perception. It is the individual who is *pulling* and transforming information to use it as proposition material. *Knowledge sharing* action mostly serves as a vehicle for interpersonal alignment. Here, the designer *pushes* his understanding towards the collective. Finally, *representation* action forms a hybrid between pushing and pulling; it serves both the individual and collective. First, regarding the collective, the projection of a proposition onto a certain medium externalizes it and as such enables others to observe it. Second, regarding the individual, the process of projection can be viewed as mapping between a set of internal and external dimensions, where a slight discontinuity exists between the two since the internal set is based on imperfect knowledge. As such, the process of projecting can present the individual with new insights.

It is important to note, that these design actions can be performed from the perspective of both the designer and the LLM-based Augmentation System (AS). The visual representation presents (figure 19) these design actions from the perspective of the designer, oriented towards the collective, of which the LLM is a part. If, for example, *knowledge sharing* action is considered from the perspective of the CA, it would represent the behavior that increases common ground and that as such satisfies part of the collaboration goals of the CA.

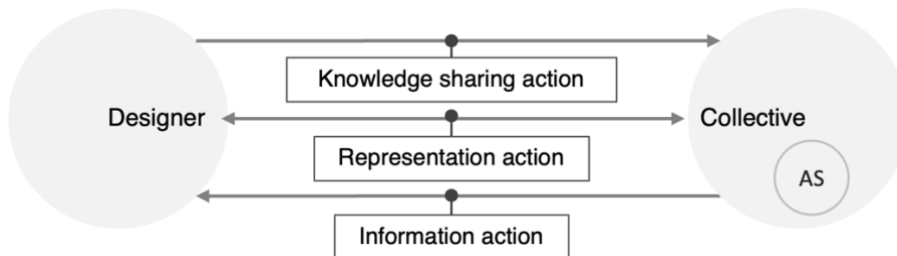


Figure 19: The purpose of the three different design actions

Based on these definitions, it can be said that *information* action is undertaken to attend to the *design* and *learning-oriented activity* layers – after all, it bridges the spaces of concepts and knowledge – whereas *knowledge sharing* action is directed towards the *collaboration* activity. Representation action forms a hybrid between the two, attending to both the design and learning, as well as the collaboration goals. The relation between the types of design activity and actions is visualized in figure 20, using an imaginative set of goals, tasks, and actions.

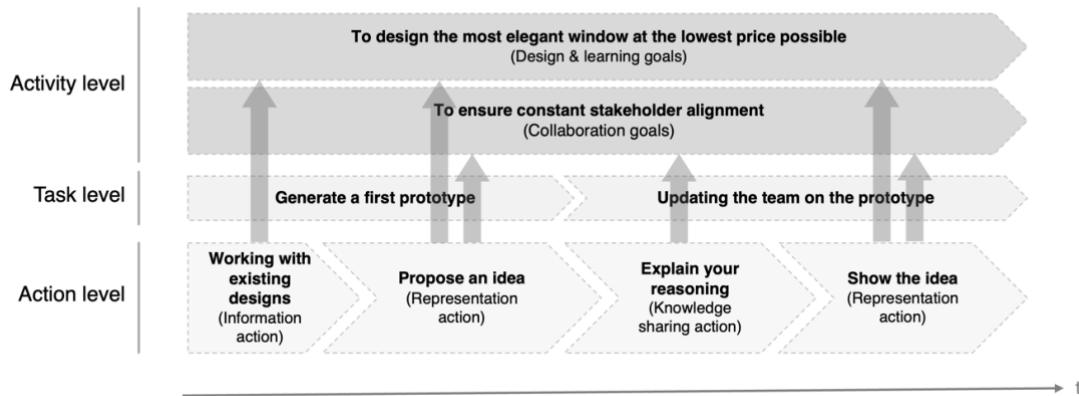


Figure 20: Design actions relative to design, learning and collaboration goals

### 5.13 Relating the layers of behavior to the unified design process framework

Uniting the findings as presented in the previous sections with the design process framework from figure 13, figure 21 shows the relation between a designer's goals, actions and the design process that unites them. It also proposes a more detailed description of the function of meta-cognitive processes that wasn't represented in earlier works. Previously, meta-cognitive processes were described as the processes that identified the limits of one's understanding. However, there was no explanation of how the content-related direction of the next action was constituted and as such how an individual is able to know in which direction to go to further his understanding. For example, if one identifies that one cannot explain the workings of a window-locking mechanism, what frame of reference then steers one's next action to alleviate such a short-coming? In other words, with regards to what desired target state should one start to feel uncertain about this inability to understand the locking mechanism?

Building on the FBS framework, one could argue that when a *structure* lacks the resolution required for explaining how a definition of *function* can be transformed to one of *behavior*, one might get an increased level of epistemic uncertainty. However, even then, the question is what defines whether the designer reformulates the description of *function*, *structure*, or *behavior* in the case where  $B_e^i$  and  $B_i$  do not correspond. In other words, what defines the *direction* of the next action? The visual presented in figure 21 is able to describe this target, as the target is represented by the goals in the activity layer and subsequent tasks that attend to these goals. It shows that perceived uncertainty originates from a certain level of tension, identified through meta-cognitive processes, between one's (in)ability to make an interpretation based on the (in)availability of schemata, and a task that is a derivative of the desired target state and corresponding direction. It thus describes a frame of reference (*tasks* and *goals*) against which the success of current *actions* can be measured through meta-cognitive processes. Moreover, it also informs our view on the ability of such tasks and goals to inform the next type of action and as such the direction of the design process. Finally, it also underwrites the claim that *learning* and *design* goals are aligned (Viros-i-Martin &

Selva, 2021). First, one cannot explain a certain phenomenon. Second, given the desired design target (design goal) one has to learn how to interpret that phenomenon as it is related to the desired design outcome. Third, one perceives a certain level of epistemic uncertainty and as such aims to work towards a better understanding by learning how to explain the phenomenon. This simple example shows how learning goals can be regarded as derivatives of design goals.

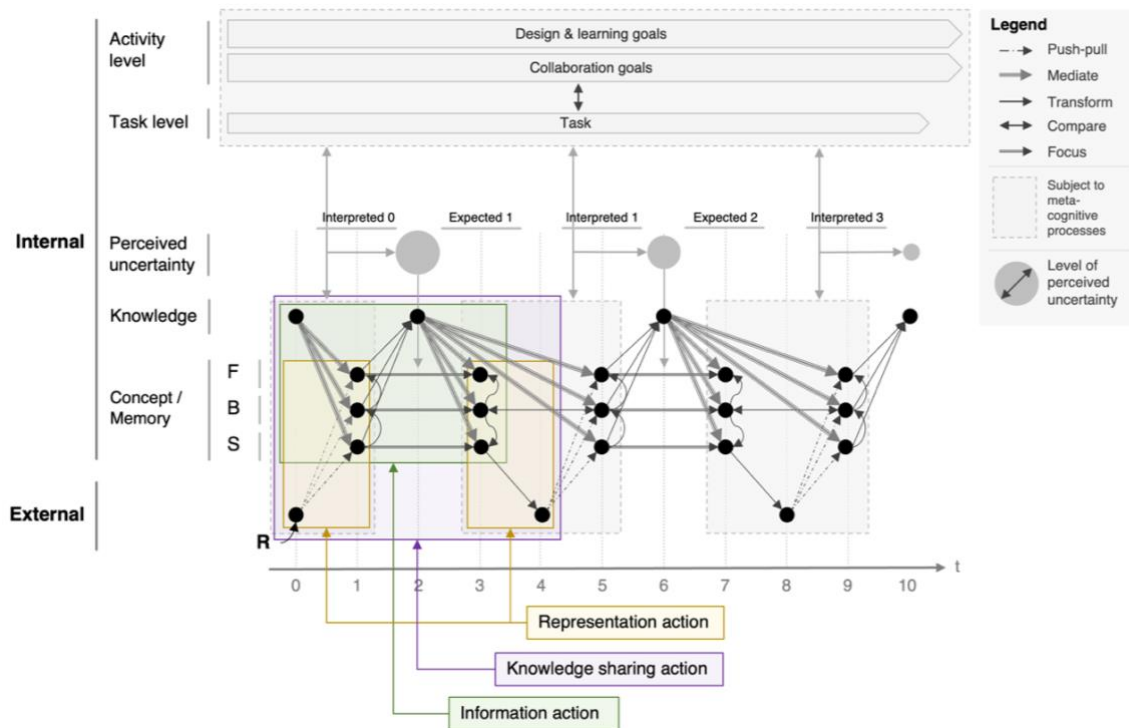


Figure 21: Layers of behavior relative to the design process framework

### Takeaway box 5

The UDA model of design describes how design behavior is the outcome of the interaction between one's motivations and goals (activity level), tasks that attend to these goals (task level), and sequences of actions (representation, information, knowledge sharing) that attend to certain tasks (action level). Each type of action relates differently to the design process framework. Representation action is the translation from an internal to an external representation and as such covers the transition between the space of concepts and the external world. Information action is the part where the designer is puzzling with the presented information and data parts given the knowledge that he has at his disposal. Finally, knowledge sharing action is the reflective practice of describing the line of reasoning behind a certain proposition. The presence of goals in the activity level also helps to describe the why behind one's perceived uncertainty, as this concept can now be described as emerging from a tension between one's (in)ability to make an interpretation and a task that is a derivative of the desired target state (the goal).

### 5.14 From design actions to mental iteration stages

This section aims to further the integration between design actions as described by Cash & Kreye (2017) and the design process framework that represents the unification of three theories of design, to build towards a more elegant (mutually exclusive) definition of the cognitive activity of a designer and his points of interaction with the external world. First, a reflection on the conceptual effectiveness of the design actions as proposed by Cash & Kreye (2017) is given to identify the limitations of their categorization of design actions for the purpose of the design process framework conceived for the purpose of defining design process augmentation. Second, a consolidation of the current theoretical findings will be provided leading to the identification of three distinct stages that describe a mental iteration of the design process. Third, these stages will be connected to design actions to understand how different types of actions relate to the different stages of a mental iteration.

#### 5.14.1 Limitations of design actions as basis for augmentation

Linking back to the question posed at the beginning of Part 2 of this section (recall section 5.11: “the question then becomes what the relation is between the design process framework as conceived so far and the actions that result from it”), the framework as visualized in figure 21 has provided a view on the relation between design actions, their origins in the activity and tasks layers, and the spaces involved in design progression (concept/memory & knowledge). However, they do not yet form an intuitive and exclusive basis for the definition of types of augmentation. As previously mentioned, in collaboration the design actions often come in pairs of two or three actions. Recalling from the previous section, every design action of a designer is either a combination of 1) *information* and *representation* action, 2) *knowledge* and *representation* action, or 3) a combination of all three actions. Since *representation* action is present in every type of design action pair, it can be stated that the augmentation system gets presented instances of information and/or knowledge sharing action packaged into a certain representation. This representation can be language, images, drawings, or any other medium.

Given this observation, one could argue that design actions are not the most elegant set of components to use for the formulation of augmentation practices. Recalling the working hypothesis of this second part of section five (the point where both actors are able to perceive and thus interact with each other is where an actor externalizes the outcome of its cognitive processes into an action that is then received by the other actor(s)), it is important for such actions to be clearly defined, for the other actor to be able to augment. As such, a mutually exclusive definition of a designer’s actions is important as the augmentation system has to correctly interpret the input it receives. The next section therefore aims to build on the definition of design action and to reformulate them towards design process stages that can be augmented. The approach taken here is to consolidate the findings from previous sections about the

design process framework and to link them to the design actions to come to a more mutually exclusive frame on design process action.

#### **5.14.2 Consolidation towards three stages**

As stated in the takeaway box 4 in section 5.9, Problem-Solution (P-S) co-evolution is a descriptive derivative of the co-evolution of the *Knowledge* and *Concept* spaces. In other words, P-S co-evolution depends on co-evolution at the level of Knowledge-Concept. Within the *Concept* space, the three stages *proposition* (design state space), *interpretation* (interpreted space), and evaluation are identified. *Proposition*, being the design state space (Gero & Kannengiesser, 2017), and *interpretation*, being the space of interpretation, can be regarded as what Crilly (2021b) describes as ‘nested’ co-evolving spaces: the co-evolution of the *proposition* (P) and *interpretation* (I) spaces acts at a more detailed level within the space of *concepts* (C) and defines the state of the space of concepts. From now on, this will be abbreviated to ‘K-C(P-I)’ co-evolution. Although the spaces of Knowledge and Concepts describe on a high level the architecture that underlies one’s ability to identify either the problem or solution in an idea, the state of the space of concepts is thus defined by two co-evolving spaces at a deeper level. The process of evaluation between interpreted and expected world then represents the fitness-function of one’s design process by establishing an account of the fit between the expectations of the proposition’s behavior (*proposition* space) and the interpretation of the actual proposition’s behavior (*interpretation* space). The fit, or the lack of it, leads to a certain instance of perceived uncertainty, whereas the content on display relative to one’s goals defines the direction of the next action. One’s knowledge is used as input and updated in this process. The three stages are visualized in figure 22 and framed within the concept of a mental iteration. First, the three stages of a mental iteration are described below and further related to the design theories described in prior sections. Afterwards, the rationale behind the definition of the mental iteration is explained.

#### **5.14.3 Explanation of the three stages**

This summary of K-C(P-I) co-evolution presents three distinct actions that partially overlap with the design actions as derived from literature by Cash & Kreye (2017): *proposition*, *interpretation*, and *evaluation* (figure 22). These three stages describe how one can progress through a process of design, or propositional learning, based on the recurring and iterative process of proposing, interpreting, and evaluating design artifacts. These three stages also align with the stages of *variation* and *selection* as described by the concept of evolutionary creativity (Simonton, 1999a, 1999b), the theory underlying the framework of AI-support for stimulation in design ideation from (Berni et al., 2024). This concept of evolutionary creativity is particularly relevant as it can be applied to any form of system (including AI). In contrast, most other creativity theories originate from the field of Psychology, and as such only apply to human behavior (Thoring et al., 2023).

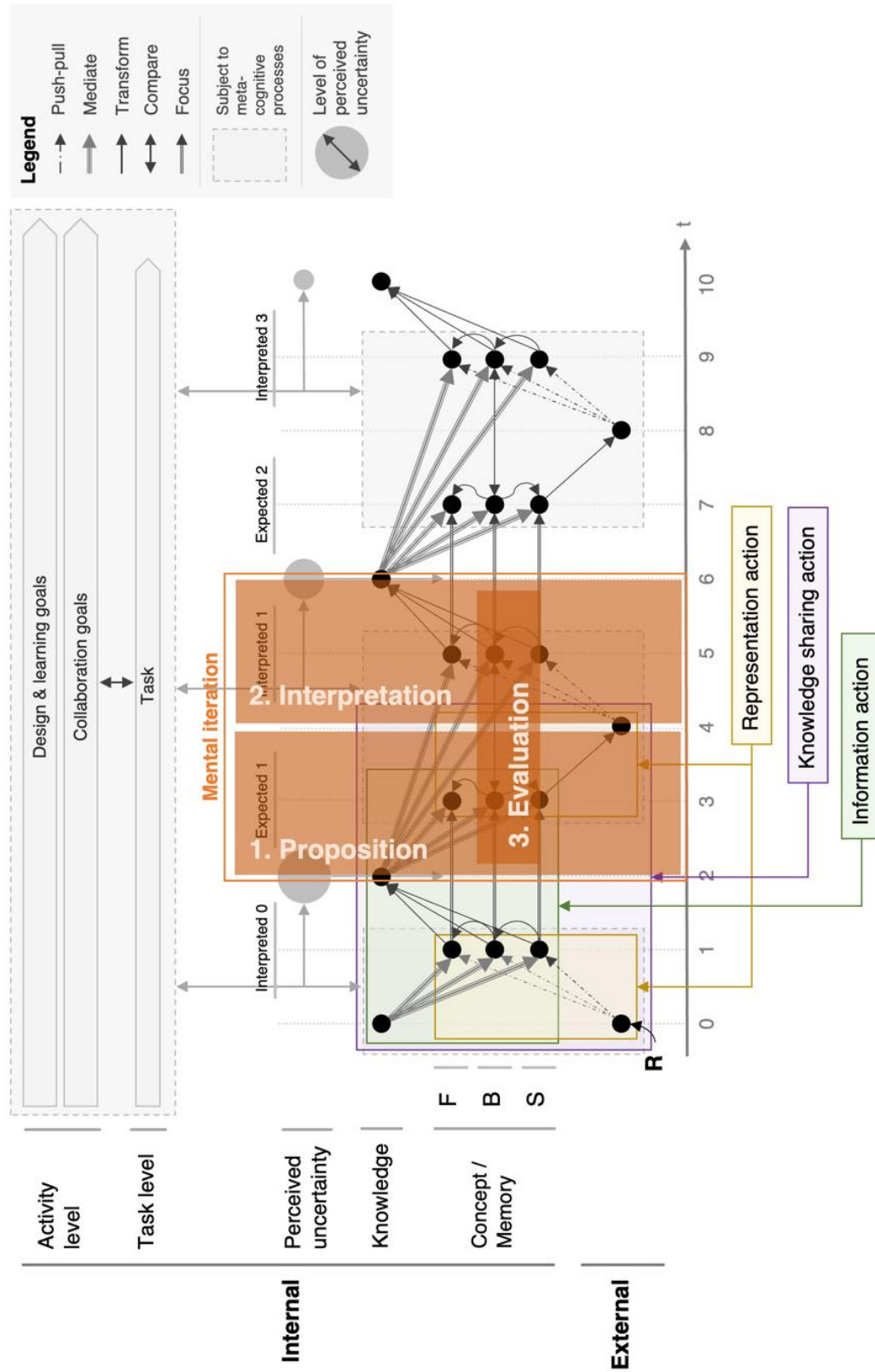


Figure 22: Three-stage mental iteration based on the design process framework



First, *variation* describes the mechanism that is used while making a proposition, comprising both *mutation* and *recombination*. Other takes on these two forms of variation include the forms of *combinational*, *exploratory*, and *transformational* creativity proposed by Boden (2009). Here *combinational* creativity can be equated to *recombination*, *exploratory* to *mutation*, and *transformational* to a *combination of mutation and recombination* (see section 6.5 for a detailed discussion). Second, *selection* can be viewed as the successive step to the stages of *interpretation* and *evaluation*. *Selection* uses the rationale resulting from the *evaluation* to update one's design state space – or expected world (see situated FBS framework) – and as such to only save those propositions that align with the latest version of the 'fitness-function'.

### *Proposition*

First, the designer aims to fill the identified gaps in the causal relationships through partial propositions of *structures* that represent the premise for the desired state: the effect or *behavior* of the final design solution. When this proposition is a one that aims to solve the current state of the design problem definition, this stage can be equated to the ideation phase in design (Berni et al., 2024). When the proposition is an individual's act of pursuing a certain research direction to understand and/or (re)define the design problem, it can be regarded as part of the problem identification/definition phase. Both forms of proposition, or 'stimuli', help the designer in constructing both the spaces of problems and solutions, by mapping between relations, functions, and paradigms (Berni et al., 2024). For example, the proposition of an idea that aims to solve the problem, may point at the existence of another element of the problem. Similarly, the redefinition of the design problem may infer another idea about a solution. Although this sounds familiar to descriptions about problem-solution co-evolution, one has to remember that before one can identify the problem, one first has to make the *interpretation* (nested in the concept space), as such updating the space of concepts, followed by an *evaluation* that compares the fit between the *interpretation* with prior *expectations* underlying the externalized artifact. Only if there is no fit, the idea, or proposition may be regarded as a problem. In the words of the addressed design theories, this stage can be described as follows.

- Problem-solution co-evolution: what is proposed is initially perceived as being the solution to the problem, that is being proposed due to a perceived lack of understanding. Propositions can be either problems or solutions depending on the level of understanding of an individual.
- Situated FBS framework: Propositions can be conceived as instances of structure that lead to the desired behavior situated in the 'expected world'.
- Concept-Knowledge design theory: Propositions are the ways in which an individual partitions its object of investigation such that it expands the meaning of it. Propositions are therefore instances of expansive partitions.
- Schemata theory: Propositions can be understood in terms of schemata as the composition and instantiation of subschemata (given a dominating schema) that lead to a certain state of the dominating schema's variable(s).

With regards to the connection between this stage and the design actions as described in the Uncertainty Driven Action (UDA) model, the proposition stage can be described as a combination *information* action and *representation* action, driven by an uncertainty perception. Proposition mostly overlaps with the second part of the definition of *information* action, defined as a combination of working with and manipulating data. Combining this observation with the representation action, the stage of proposition can thus be described as manipulating data with the manipulation being externalized through the projection of it onto a certain medium.

### *Interpretation*

Second, the designer has to identify the types of knowledge components (domains) and their causal relationships to interpret the proposition, or stimuli. The dimensions as presented and triggered by the object (e.g. 'R') challenge the individual to construct a representation based on his/her existing knowledge and memory. One could argue that this is the process of constructing an *embedding* of the presented environment. Often this leads to the realization that the status quo of knowledge is not sufficient for defining the embedding at the desired level of detail, due to gaps in the chain of causally related knowledge components. As will become apparent in the next stage of evaluation, the embedding is dynamic as the designer learns new components and relations throughout the process. The topic of 'design fixation' is especially relevant in this stage, as a lack of openness or an imperfect conviction that that A leads to B might lead the designer to construct an interpretation that is not true, relevant, or appropriate given the contextual conditions; the designer might become blind to other interpretations. In the words of the addressed design theories, this stage can be described as follows.

- Problem-solution co-evolution: the object of understanding represents either a problem or solution to the designer. When a designer is capable of constructing a perfect set of propositions to explain its desired behavior, it is considered a solution. When not, it is regarded as a problem.
- Situated FBS framework: Understanding can be equated to constructing the interpreted world based on push-pull processes between an external world, experiences and memories.
- Concept-Knowledge design theory: Understanding is one's ability to identify the right set of knowledge 'partitions' that describe the phenomenon of investigation.
- Schemata theory: Understanding is the process of the selection of schemata and the construction of their composition based on memories of previous schemata compositions.

Relating this stage to the design actions from the UDA model, it can be defined as the inverse of proposition: a combination of *representation* and *information* action. Instead of projecting one's thoughts on a medium, the instantiated medium has to be interpreted and translated to a set of instantiated schemata. This presents an interesting

observation. Whereas a paper is an external form of a medium, one's knowledge structure can be regarded as an internal medium. Both the drawing and one's memory are the result of expressing a certain set of datapoints onto that medium. Putting aside this observation, it can thus be said that *interpretation* is the act of extracting the information captured in an external representation and embedding that information in an internal medium through *information* action (working with data).

### *Evaluation*

Third, the designer compares the interpretation of the proposition to the expectation about *structure* that preceded the externalization of it. The dimensions along which the interpretation is conceived are compared to the instantiated set of dimensions as represented by memory to identify similarities or discontinuities. In the words of the addressed design theories, this stage can be described as follows.

- Problem-solution co-evolution: Evaluation can either lead to the observation that the proposition represents a solution, as expectation and interpretation align, or that the proposition presents previously unattended dimensions, leading to the conclusion that it presents a new problem.
- Situated FBS framework: The stage 'evaluation' can be equated to the 'evaluation' stage in the situated FBS framework, describing the process of comparing between  $B_e^i$  and  $B_i$ .
- Concept-Knowledge design theory: Although the C-K theory doesn't make a distinction between different stages, it can be said that evaluation is one's act of comparing the current set of knowledge partitions to past instantiations of it.
- Schemata theory: Evaluation is the process of comparing between two versions of instantiated schemata to identify the similarities or differences in the composition of subschemata and variable values.

Linking back to the UDA model, one could argue that there is a loose connection to be drawn between *evaluation* stage and *knowledge sharing* action because they both operate at a higher level as they are able to observe the actions that have been taken. Whereas *knowledge sharing* action is the act of explaining the reasoning behind one's actions to others, *evaluation* is the act of comparing a description of one's reasoning before and after the externalization of the artifact. It is about comparing the internal embedding of the artifact (the *interpretation*) to the expectations underlying the manipulation of a set of datapoints that lead to the creation of the artifact (the *proposition*).

The previous paragraph also implicitly explains the notion of *emergence*, defined as the process of uncovering something existing, as well as being presented with something new (Dorst, 2019). For example, one might choose to buy a pair of jeans in size medium from brand A because one has bought jeans from that brand before that fitted well. When fitting the new pair of jeans, it stands out that the jeans do not perfectly fit as they are a slightly different model. Hence, there is a difference between

expected effect, a perfect fit, and reality, a non-perfect fit. Why? Because the individual wasn't aware of the 'dimension' of jeans-types that influences the way the pair of jeans fit one's waist. Based on this observation, the individual from now on associates the schemata of jeans-types to the process of buying jeans. Based on this updated set of schemata and the instantiated schemata that connects it to the dominating schema of 'buying jeans', the individual continues its search for propositions that satisfy the desired behavior. First an individual, driven by perceived uncertainty, wants to better understand the context by proposing a *structure* (or idea) that aims to describe a perceived reality (uncovering something existing). Second the individual tries to embed the *structure's* behavior into the schemata that was used as the basis for the idea (interpretation), only to identify during *evaluation* that the actual behavior differs from that described by the current set of instantiated schemata (being presented with something new). The individual recognizes the need to search for other schemata that are required to arrive at the desired *behavior*. This simple example shows how the process of design is not exclusive to big design projects, but also occurs in many almost unconscious moments during one's life. One could even argue that what is referred to as design is the most fundamental process that allows humans to learn and grow beyond existing knowledge by further exploiting the affordances presented by our material and social environments.

## 5.15 Types and scales of mental iteration

### 5.15.1 Global and local mental iterations

The design process is also often described in terms of consecutive phases, where each phase has a different focus or purpose. Moreover, such descriptions are often described at various scales. For example, approaching the design process from the scale of the entire design project, the double diamond framework describes a process of first mapping the problem space to arrive at an accurate problem definition, to then start generating ideas to solve the problem<sup>10</sup>. Here, where each phase is a combination of diverging and converging movements. Another design process description that acts at the level of a mental iteration describes a four-phase process: *analyze* (problem), *generate* (idea), *compose* (concept), and *evaluate* (concept) (Zhou et al., 2023). Here, mental iterations are defined as a “*cycles connecting between forms of logic*” and “*bridging the gap between information processing and decision-making*” (Zhou et al., 2023, p. 3,4). They can occur locally, at the level of an individual phase, or globally, in between and as such interwoven throughout the four phases (Zhou et al., 2023). For example, a local mental iteration can be the consecutive iterations made within the phase of problem analysis, or idea generation. A global mental iteration is a complete string of instances of problem analysis, idea generation, concept composition and

---

<sup>10</sup> From an [article](#) (2005) at the design council's website, named “Framework for Innovation: Design Council's evolved Double Diamond”.

evaluation (Zhou et al., 2023). Combining this logic with the design process framework from figure 22 it can be said that each set of the three stages is a mental iteration at the local level. As already described above, such a cycle of three stages can be used to describe how one generates an idea, or chooses a research direction, and as such forms the basis for describing various the four types of local mental iteration presented by Zhou et al. (2023). Given this classification, the design process framework is thus able to express different phases in the design process. Four such local mental iterations then describe a global mental iteration. This relation between these types of mental iterations, their scale and the design process framework is visualized in figure 23.

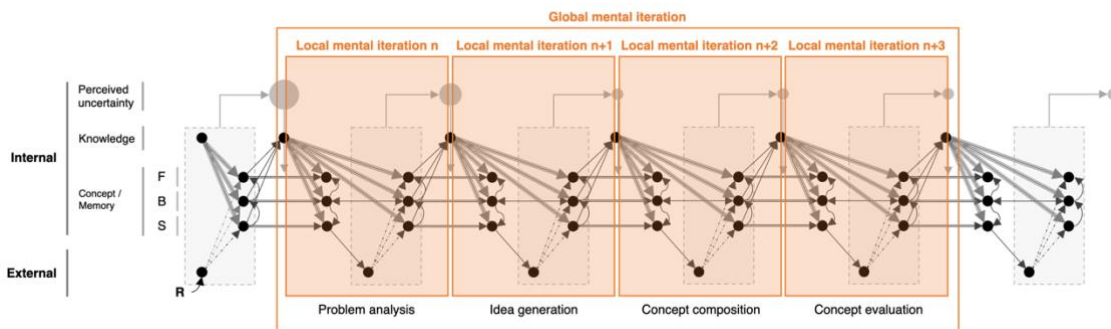


Figure 23: Design phases and mental iteration scales

#### 5.15.2 The function of global and local mental iterations

The previous paragraph focused on describing mental iterations at the scale of expanding one's understanding through the *proposition*, *interpretation*, and *evaluation* of *structures*, or simply *ideas*. However, a design *concept*, as conceived in the concept proposition stage (figure 22), differs from a single *idea* proposition. It is an idea about the composition of various *ideas*, or in the terms of the FBS framework, an idea about *structure*, originating from the individual's account of knowledge relative to the design brief. Expressed in terms of schemata, *ideas* can be represented by subschemata and their subschemata and corresponding variable values based on variable distributions. *Concepts* are the integrated *structure* of subschemata (or *ideas*) that aim to satisfy the desired *behavior* of the dominating schema. How then is the origination of such *concepts* described by the design process framework at the level of mental iterations? One could argue that, whereas a local mental iteration describes the progression of *ideas*, a global mental iteration describes the progression of *concepts*. In other words, ideas are vehicles for concept progression. Below, four local mental iterations (LMIs) are described that together form a global mental iteration (GMI).

##### - Problem Analysis phase

First, the designer has a certain conceptual understanding of the problem space that, relative to the design goal, does not have the desired resolution. Expressed in terms of the FBS framework, the set of functions retrieved from one's memory does not yet satisfy the final goal of the design brief. This leads the designer to propose a function, or phrased differently, an *idea* about the problem.

- **Idea generation**

This proposed *idea* about the problem, once interpreted and embedded in the designer's conceptual understanding space, leads to the proposition of an *idea* about a solution – a part of the *structure*.

- **Concept composition**

After interpreting the idea about a possible solution and relating it to one's understanding, a proposition in the form of a *concept* is made where, instead of introducing a new component, an overview of one's account of the design's *structure* is presented.

- **Concept evaluation**

Finally, both the individual *ideas* about either the problem or solution, or the idea about structure (the *concept*) may be reformulated as a result of one's evaluation of the behavior of those propositions relative to one's expectations.

#### Takeaway box 6

Since the design actions as presented in the UDA model of design weren't mutually exclusive once mapped onto the design process framework, a new description of a designer's actions has been conceived that does form a clear basis for augmentation practices. These newly conceived action definitions (proposition, interpretation, evaluation) are formulated as stages of a (local) mental iteration (figure 22). This mental iteration represents a certain form of design activity (e.g. problem analysis, idea generation) and represents design progression at the level of ideas. A sequence of local mental iterations (LMIs) forms a global mental iteration (GMI). Whereas LMIs describe design progression at the scale of ideas, GMIs describe design progression at the scale of concepts.

### 5.16 A definition of the human actor in the design process

Based on the conception of the design process as visualized in figure 22, each category of the human-AI collaborative design space exploration framework (Viros-i-Martin & Selva, 2021) can be detailed for the human actor (table 7). For example, the driver of the design process – the spaces of knowledge and memory/concepts in combination with epistemic perceived uncertainty resulting from meta-cognitive processes – is represented by the categories of *knowledge* and *internal state* respectively (table 7). Next to that, the three types of goals described in the original framework can be described as being part of the activity layer from the framework of designer behavior (Cash & Kreye, 2017). Consolidating these observations from the perspective of the human actor, the following can be said: Initiated by the designer's perception that comprises his interpretation – or the understanding stage of mental iteration – of the design context and inferred state of the augmentation system, supported by his spaces of memory and knowledge, driven by perceived epistemic uncertainty, steered by his goals, and constrained by his attributes, the designer takes action.

Human-AI collaborative design space exploration framework				Mental iteration design process
Perception	Estimated state of the world	Human	An approximation of the state of the design task	Part of the interpretation stage, relating to the <b>concept/memory space</b> stage of a mental iteration.
	Inferred state of the CA	Human	The expected interpreted state of the CA given the interaction	[Interpretation of the state of the LLM]
Factors that constitute the state of the agents	Internal state (Dynamic during design process)	Human	Motivation, cognitive workload, situational awareness, <i>trust in CA</i> , <u>design preferences</u> , the <u>degree of design fixation</u>	Defined by the level of perceived <b>epistemic uncertainty</b> and as such dependent on the design task <b>relative to the designer's knowledge</b>
	Knowledge	Human	Domain, design process, problem specific	Relating to <b>knowledge space</b> in the design process framework, which is <b>dynamic</b> and extendable
	Goals	Human	Design, learning, collaboration	A designer's goals are positioned within the 'activity' layer of a designer's behavior. The <b>learning</b> goal is to alleviate one's perceived uncertainty through the proposition of candidate ideas. This is aligned with the design goal. The <b>design goal</b> follows the same mechanism [collaboration goals]
	Attributes (Static during design task)	Human	Cognitive style, level of expertise, risk aversion, expectations, preconceptions	[Person-based attributes]
Legend <ul style="list-style-type: none"> <li>- <i>Italic</i> = Related to human machine collaboration literature*</li> <li>- <u>Underlined</u> = Related to design literature*</li> <li>- <b>Orange</b> = Not covered so far by prior literature in this thesis.</li> </ul> *(Viros-i-Martin & Selva, 2021)				

Table 7: The design process related to the framework for human-AI collaborative design space exploration.

Table 7 also shows the areas that are not yet defined. For example, it is not yet described how the designer is able to infer the state of the CA, or augmentation system. The next section will further detail this point. Moreover, collaboration goals are not yet detailed. According to Viros-i-Martin and Selva (2021), collaboration goals inform how to approach the interaction with the other agent, where the approach is often composed by a designer's expectations and preconceptions with respect to the capabilities of the CA agent. As such this depends on both the general familiarity of the public with the capabilities and limitations of the augmentation system, and the specific individual's knowledge about augmentation system. As such, it seems that for proper collaboration to emerge, there should be upfront alignment on the preconceptions about and expectations of the collaborative act. *Expectations* and *preconceptions* are also stated within the category of person-based attributes. Similar to *expectations* and *preconceptions*, for augmentation to succeed, it is also important to align on one's expertise. Furthermore, it might be necessary to share one's level of risk aversion and cognitive style, although the specific details go beyond the scope of this thesis.



## 6. Design process augmentation

---

Augmentation comes in many forms. Engelbart (1962) already provides the context for these forms, arguing that augmentation happens in the context of *learning, language, methods, and artifacts*. This section is specifically interested in the augmentation of the design process as conceived in the previous section. It therefore attends to the third research question:

*How can the design process be augmented?*

---

### 6.1 Approach

To answer this question, the definition of human intellect augmentation as defined in section three should be connected to the three stages of the design process framework (as visualized in figure 22). This leads to formulation of the three ways in which the design process can be augmented at the level of a mental iteration, called the three types of design process augmentation. Afterwards, a connection between the definition of a successfully augmented process and the definition of a creative product, as described in creativity literature, is observed. This observation of the relevancy and fit of creativity literature in the domain of design process augmentation not only allows for more a more detailed formulation of the desired characteristics of the outputs of an augmented process, but also enables the translation of that formulation to one that describes the process-related attributes that lead to such outcome-based augmentation success. This results in six forms of design process augmentation, where each of the three types of augmentation can focus on two dimensions: *novelty* and *quality*. To make the step towards the practical manifestation of these descriptions of augmentation forms, they are embedded in both different types of *personas* as well as *interaction views*. These two embedded forms of the six forms of design process augmentation can be integrated in between a designer's mental iterations to form a string of sequentially ordered human-AS (Augmentation System) collaborative design actions. This creates a conceptual framework of human-AS co-design that forms the basis for answering the research question of this section.

### 6.2 Defining design process augmentation at the level of mental iterations

Recalling from the beginning of this thesis, design process augmentation can be defined as increasing the capacity of a designer to approach a complex problem situation, to gain comprehension, and to derive solutions to problems in a design project (adapted

from Engelbart, 1962). This definition can be linked to the layers of design activity and the three stages of a mental iteration (as in figure 24).

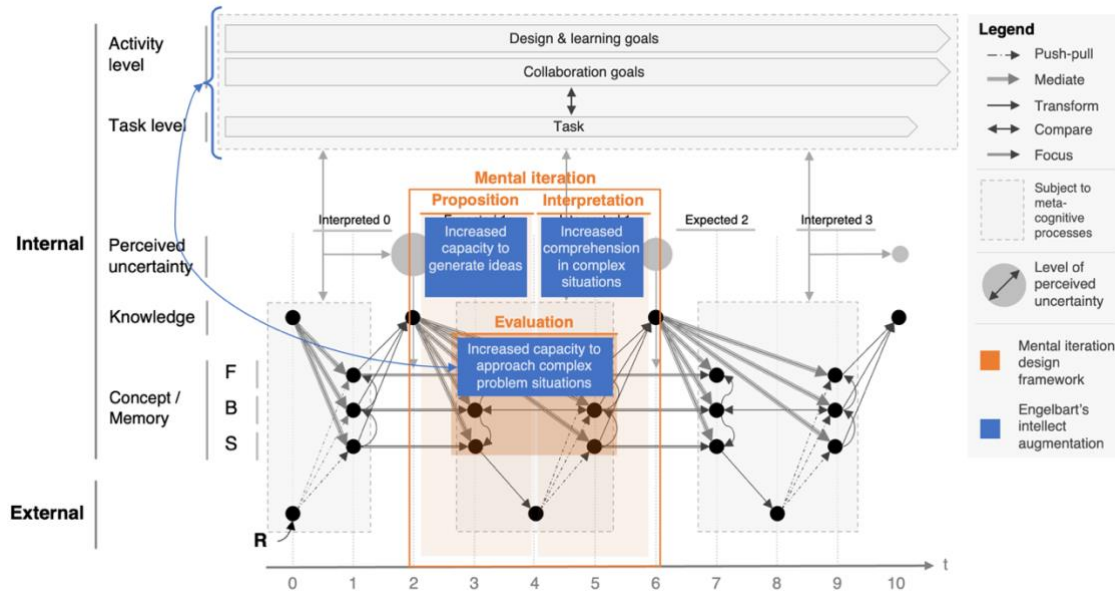


Figure 24: Engelbart's augmentation related to the design process framework

First, 'increasing the capacity of a designer to approach a complex problem situation' can be equated to supporting the designer in constructing the right set of goals and tasks to which future actions attend; to establish an account of direction. However, establishing this direction once is not enough. Building on the co-agency framework (Glăveanu, 2015), that describes a person's ability to act in the world as a cyclical process of *intentionality* (one's general orientation towards the world based on the aggregate of one's goals), (re)action (here, one can mentally read *information, representation, knowledge sharing* action (Cash & Kreye, 2017)), and reflexivity (the ability to change perspective on one's action), it can be argued that the stage of *evaluation* is a critical aspect in ensuring the accurateness and relevancy of the definition of one's goals. As such, it is more accurate to equate this part of the augmentation definition to the designer's ability to *evaluate* on the effectiveness of an action relative to the task and goal it attends to. This then leads to either a reformulation of the task and goal, or of the type of next action, based on a progressed understanding after the proposition and interpretation stages.

Second, the 'to gain comprehension in complex situations' can be equated to supporting the designer in its *interpretation* stage and as such with *information* action, as 'comprehension' can be regarded as a designer's ability to interpret the situation given the set of causally structured knowledge components he or she has at disposal (see section 5.7). It is the act of forming a conceptual understanding of the design context through the interpretation of new and prior propositions in the form of artifacts and memories respectively.

Third, ‘increasing the capacity to derive solutions to problems’ can be compared to supporting the *proposition* stage of the mental iteration. This part of the definition will be reformulated to ‘increasing the capacity to generate ideas’ to make the division of Engelbart mutually exclusive. This has to do with the definitions of ideas and solutions. As concluded in the previous section, an idea is a proposition that, after having externalized it and interpreted it, is embedded in one’s conceptual understanding. A solution is defined in this thesis as the formulation of one’s conceptual understanding such that it explains how a perceived problem is solved. That is, it connects desired *function* to *final behavior* via a composition of *structure*. Generating solutions is thus the combination of proposing ideas (*proposition*), integrating them into a larger web of knowledge (*understanding*) and additionally formulating that *understanding* such that it forms a coherent story between problem and solution.

Based on this categorization, one can argue that a designer can be augmented in three ways, referred to as bottom-up, top-down, and directional augmentation (table 8). These three types of design process augmentation translate the factors related to human intellect augmentation into the language of the mental iteration design process framework and are detailed below.

Mental iteration design process framework	Human intellect augmentation	Augmentation types
Proposition	Increased capacity to generate ideas	Top-down
Understanding	Increased comprehension of complex problem situations	Bottom-up
Evaluation	Increased capacity to approach complex problem situations	Directional

Table 8: Relation between mental iteration design model & human intellect augmentation

These types of augmentation can be understood in the light of the following analogy. Imagine being the driver of a car. The driver can be assisted in defining and evaluating the direction in which he should be or is driving (directional). He can also be supported in reading the road signs or traffic situations that occur along the way (bottom-up). Finally, the car may also intervene by automatically braking or adjusting the steering angle to stay in line if the driver does not operate punctually (top-down). Below a more detailed explanation is provided for each type of augmentation.

#### 1. Bottom-up augmentation

The first type is called bottom-up augmentation, as it reinforces the basis for the *proposition* stage. Here a designer is supported in gaining a better understanding of the situation through additional information that strengthens his ability to interpret the object of investigation; the design brief or any iteration that follows

up on it. This idea is in line with the foundation theory which states that creativity originates from the cognitive work performed on top of past knowledge (Weisberg, 1999), where the augmentation support increases the basis of past knowledge. Knowledge is defined as the separate information components that one possesses. However, during interpretation, the knowledge is not enough for understanding a situation. For this, the designer needs to construct an explanation of how these information components relate, and hence what their causal structure is both relative to each other as well as to the design problem. As such, a designer should not only be augmented in its basis of knowledge of factual issues, but also its knowledge of causes – its *understanding*. For example, in the window-design example (Gero, 1990), the designer can be augmented by providing him with more relevant contextual information about the characteristics of the wall in which it will be placed, the indoor room, and the outdoor environment that surrounds it. Here the ‘knowledge of causes’ would explain how various values of the ‘variables’ (contextual information) would affect the overall outcome. The existence of, for example, a large tree next to the place where the window would be located might lead the designer to increase the dimensions of the window such that sufficient light reaches the indoor interior. Moreover, the wall thickness might provide the designer with cues about the level of shade that falls over the window and hence also the reduction in the amount of light entering the room.

## 2. **Top-down augmentation**

Although the designer might be able to come up with a set of ideas, this will always be a subset of all possible solutions, as it is based on incomplete knowledge. Therefore, a designer can be supported in generating propositions, structures of expected causally related logical knowledge components; ideas that aim to solve the apparent disconnection between the status quo and a desired state. This type of augmentation is referred to as top-down, as it inspires the designer, pulling him in several potential solution directions. Here, the designer is provided with propositions to choose from. Before any example can be considered valuable, it needs to be understood first. It is therefore critical that the designer is capable of interpreting and understanding the provided example. As such, a hybrid of bottom-up and top-down can be conceived as an additional type of design process augmentation.

## 3. **Directional augmentation**

This type of augmentation is intended to assist designers in choosing a design direction that leads to optimal results. Rather than focusing on directly enhancing the content itself, it is a more indirect approach towards improving the output, as it focusses on improving reflection on the content in relation to one’s goals by augmenting one’s evaluation process. This means, that the system proposes information that may affect the explicit or tacit (re)formulation of one’s

goals and tasks. This approach helps to clarify the factors that influence design direction, supporting designers in aligning their goals and actions, especially when new information emerges that affects their knowledge and thus reshapes their understanding of the design space and desired goal, task, or action.

### 6.3 An outcome-based definition of a successfully augmented process

Given these types of augmentation at the level of a mental iteration, the question arises what a successfully augmented process looks like. The word ‘better’ in the paraphrased definition of intellect augmentation from Engelbart – allowing a designer to *better* solve complex problems – implies the following characteristics: ‘speedier’, ‘of higher quality’, and ‘enabling what was not possible before’ (Engelbart, 1962). Underlying these terms are dimensions that say something about the characteristics of the ‘product’ of the design process; they can be considered dimensions along which the success of the outcome of the LLM-augmented process can be measured. The relation between the definition of augmentation and its success factors is visualized in table 9. Although the first characteristic, ‘speedier’, is easy to measure and relatively simple to lead back to the process of augmentation, the other two present more of a challenge. They are therefore indicated in red in table 9. How should ‘quality’ and ‘enabling-the-previously-impossible’ be defined, and how can these characteristics be translated back to process-related attributes? The objective of answering this question is not necessarily to measure the success of augmentation (although this is an obvious use case and topic for future research), but rather to provide a better fundament for connecting these output-oriented features to process-related attributes. Creativity literature presents a way forward, by providing better definitions and processes related to the definitions.

Augmentation successfactors → Design augmentation ↓	Speedier	Higher quality	Enabling what was not possible before
Bottom-up	Time	Process?	Process?
Top-down	Time	Process?	Process?
Directional	Time	Process?	Process?

Table 9: Types of augmentation relative to a successful outcome of an augmented process

### 6.4 Towards process-related attributes of augmentation

To better understand the relation between the design process and its product, one has to visit the creativity frameworks that have been proposed by various researchers. The creativity frameworks describe on a macro level the relations between the facets that together constitute our understanding of creativity, and can be considered the ‘meta-theory’ (Glăveanu, 2013) within which all other research on creativity can be positioned.

The relevance of this meta-theory is the fact that it distinguishes different perspectives on creativity and as such enables the distinction between dimensions that measure the output of a design process, and those that describe the process that led to that output in the context of ‘actors’ and their ‘environment’. Below, three closely related creativity frameworks are briefly introduced. These creativity frameworks are fitting given the context of designer augmentation for two reasons. First, creativity is seen as the epitome of human intellectual capabilities (Boden, 2009). As such, using the theory that aims to describe and explain this construct as a basis for reasoning about the potential enhancement of the human intellect seems a good fit. Second, since design and creativity are concepts that are inter-related, where design can be investigated through the lens of creativity and vice versa (Taura & Nagai, 2010), one can mentally read the word ‘design’ every time the word ‘creativity’ is mentioned.

The close connection between design and creativity can be illustrated by showing the similarity between our conception of design as *understanding* and *proposition*, and the conception of the creative process as a combination of *selective retention* and *blind variation* respectively (Runco, 2004). Although *selective retention* is not the same as *understanding*, it can be regarded as a derivative of it. Put differently, when a designer better understands the design problem, he is better able to ‘selectively retain’ which ideas or propositions are capable of leading from the status quo to the desired situation. *Blind variation* can be compared to a designer proposing a chain of object-effect combinations translating an object to its effect.

#### **6.4.1 Creativity frameworks**

The most-used framework in creativity research is that of the 4Ps of creativity (Runco, 2004) (Rhodes, 1961; Couger et al., 1993; Horn & Salvendy, 2006; Thompson & Lordan, 1999). This framework describes a set of perspectives (*person, product, process, press*) from which to approach the construct of creativity. Glăveanu criticizes the 4P framework for presenting an individualistic, static, and somewhat disjointed vision of creativity, with little effort to detail the relation between the four perspectives (2013). He therefore proposed the 5As of creativity: *actor, action, artifact, audience* and *affordances*, dividing the *press* perspective from the 4P framework into *audience* and *affordance* to make a distinction between the social and material aspects of the *press* respectively. The 5As of creativity can be regarded as the culturally embedded variant of the 4Ps of creativity, to help relate the creativity research from psychology to other disciplines interested in researching the construct of creativity (Glăveanu, 2013). As such, it is a relevant framework for this thesis that aims to use creativity literature in the context of Large Language Models (LLMs). Addressing both the 4P and 5A frameworks of creativity, the 5C creativity framework (*collective, collaboration, contributions, community, context*) proposes a more collaborative set of perspectives on creativity within the field of human-computer co-creativity, a subfield of computational creativity (Kantosalo & Takala, 2020). The *person* or *actor* perspective has been reconceived as the *collective* that consists of at least one human and one computational collaborator

(Kantosalo & Takala, 2020). Since all frameworks have their strengths and weaknesses, all three will be used to inform our view on the way in which the product-related augmentation-success-factors can be translated to process-related features. To do so, let's consider the relation between the product (or *artifact/contribution*) and process (or *action/collaboration*) as described in creativity literature.

#### **6.4.2 The creative product & its relation to design process augmentation**

Within the 4P framework, the object resulting from the creative act is called the *product*. Many researchers consider the *product* to be the proof of creativity (Kagan, 1967; Plucker et al., 2004; Ritchie, 2001; Tardiff & Sternberg, 1988). Furthermore, researchers argue that the individual's creativity cannot be exhaustively assessed without a measure of its product (Horn & Salvendy, 2006). Besides being closely related to the *person* and *press*, the *product* is also the outcome of a process; more literally, it is the 'product' of a *process*. Just like the *person* and *press* reciprocally influence each other (Glăveanu, 2013), with the *product* being a specific means for communication, *product* and *process* are also highly intertwined; the *product* is the aggregate of *process*-related insights and characteristics, whereas the *process* can be expressed in terms of the insights extracted from various iterations of the *product*. This is in line with previous observation of the design process as a process of expanding one's understanding, with the design solution being as good as a designer's understanding of its design-context. Besides, it presents a way to formulate the *product*-related success-factors of augmentation in terms of their process-related counterparts.

First, however, a general conception of the characteristics of a 'creative product' has to be established. One of the most established works on what can be considered a creative product is that from Amabile (1983) ((Sääksjärvi & Gonçalves, 2018). She specifies a creative product as something that contains elements of both *novelty* and *usefulness*, with the latter also being referred to as *quality* or *effectiveness* (Runco, 2004). Building on this synthesis, Boden adds a distinction between psychological novelty and historical *novelty*, leading to P- and H-creativity respectively. The former describes an artifact that is novel to the person who generated it, while the latter refers to something that never occurred in the entire human history (Boden, 2009).

A clear similarity between these notions of a creative product and the success-factors of intellect augmentation from Engelbart can be observed (Table 10). Novelty can be compared to 'enabling what was not possible before', whereas usefulness (or quality) can be equated to the successfactor 'of higher quality'. Below both factors are further detailed and translated to process-related insights, starting with novelty.

<b>Augmentation successfactors</b> ➡ <b>Design augmentation</b> ↓	Speedier	Higher quality = Usefulness/ Quality (creativity)	Enabling what was not possible before = Novelty (creativity)
Bottom-up	Time	Process?	Process?
Top-down	Time	Process?	Process?
Directional	Time	Process?	Process?

Table 10: Augmentation success factors related to creative product definition components

#### Takeaway box 6

The definition of human intellect augmentation as defined in section three is connected to the three stages of the design process framework (figure 24), leading to the formulation of three types of design process augmentation (top-down, bottom-up and directional). Besides these types of augmentation, the focus within each type of augmentation is defined by the definition of outcome-based augmentation success (speedier, of higher quality, enabling what was not possible before). A connection is made between these terms and the definition of a creative product (of higher quality = usefulness/quality, enabling what was not possible before = novelty). Since creativity literature connects outcome or ‘product’-related characteristics to process-related attributes, this creates the basis for identifying process characteristics that enable augmentation output success. The combination between augmentation types (3x) and focus (2x) leads to six forms design process augmentation.

### 6.5 Novelty augmentation

First of all, *novelty* and the augmentation success factor ‘enabling-the-previously-impossible’ (Engelbart, 1962) can be related by considering the former ‘the scale’ and the latter an extreme end of that scale (figure 25). Further contemplation on the definition of *novelty* strengthens the conceptual potential of this observation. Dean et al. (2006) define *novelty* as a combination of *originality* and *paradigm relatedness* (Dean et al., 2006) in line with definitions proposed by Besemer & Treffinger (1981), Jackson and Messick (1965), and Bessemer and O’Quin (1987). Here, *originality* is defined as something that is not only *rare* but also *ingenious* or *imaginative* (Dean et al., 2006). *Paradigm relatedness* (Nagasundaram and Bostrom, 1994-95) is the factor that complements *originality* by showing how the created artifact, or the creative product, relates to the status quo. It is a factor that focusses on the degree to which an idea modifies a paradigm, determined by the degree to which the idea’s elements and their relationships are altered.





Figure 25: Relation between novelty & Engelbart's successful augmentation of intellect (1962)

Building on the definition of *paradigm relatedness*, Boden (2009) and Dean et al. (2006) present complementary views on what the process looks like that enables the *actor* to arrive at ideas with varying degrees of novelty. First, Dean et al. (2006) define the four quadrants of paradigm relatedness based on how an idea's elements and their relationships are altered relative to the status quo (figure 26). In the first quadrant – this quadrant is titled 'Refine' – the idea consists of a set of elements and relations that are not different from any other existing artifact in the given paradigm. The second quadrant (Extend) describes ideas that add new elements (dotted lines) while keeping the relationships the same. The third quadrant (Redesign) works with the same set of elements but introduces new relationships (dotted lines) between those elements. Finally, the fourth quadrant (Transform) is a combination of the second and third while also extending even further. It uses new elements that can also reside without the current paradigm, while also constructing new relationships between the existing and new elements.

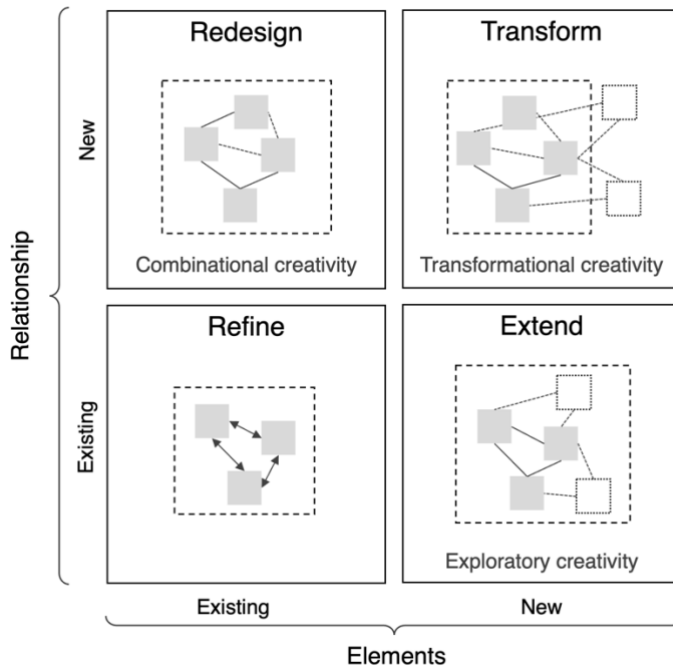


Figure 26: Paradigm relatedness relative to Boden's three types of creativity

Clear similarities can be observed between these notions of paradigm relatedness and the Boden's three types of creativity (2009) (figure 26). The 'redesign' quadrant can be compared to that of *combinational* creativity, defined as creating "*unfamiliar combinations of familiar ideas*" (Boden, 2009, p. 24), as it describes how new relations are constructed between existing elements. The 'extend' quadrant relates to the type of *exploratory* creativity, defined as exploiting the conceptual space of generative rules or 'style of thinking' (Boden, 2009), as certain relations (generative rules) are respected while new elements within the conceptual space are being explored. Finally, the 'transform' quadrant can be equated to *transformational* creativity, defined as changing the style of thinking or generative rules that define the conceptual space.

Linking back to the conception of successfully augmented process as 'enabling-the-previously-impossible' (Engelbart, 1962), the process should thus enable an individual to arrive at ideas or concepts that have never been conceived before (or resulting in 'historical novelty' (Boden, 2009)) based on highly imaginative (original) perspectives and by making unfamiliar combinations of items that are new relative to the existing paradigm. As such, the definition of an existing paradigm might be challenged, resulting in a different approach to its definition. For example, the paradigm of a car-based mobility can be seen as ultimate freedom for the individual. However, when adding previously unconnected factors such as 'cities' and the 'environment', the perspective shifts from the individual to the collective domain. This transforms the paradigm, leading to vastly different views on the desirability of the car in its current form. The following subsections paraphrase what is described above from the perspective of the three types of design process augmentation (bottom-up, top-down, directional).

#### **6.5.1 Bottom-up novelty augmentation: increasing one's understanding to fuel novelty**

The distinction between *elements* and their *relationships* aligns nicely with the two components underlying bottom-up augmentation: knowledge components and their causal relations. Bottom-up augmentation can therefore be conceived as

1. introducing new knowledge elements (schemata), that, if applicable, are positioned remotely from the initial domain; that is, they originate from other domains of knowledge, expertise, or other forms of categorizations.
2. presenting new ways in which these knowledge components can be related to extend existing ways of looking at the design challenge.

Using the window-design example again, conceptual comparisons between the design of a window and the opening in a dam can be made. This could be considered a form of exploratory creativity as, potentially, new elements present in the dam can be introduced in the window. It is, however, not yet transformational as it does not change the relationship of a) the window components, or b) the window relative to the rest of the structure of the house. To realize transformational novelty augmentation, domains of knowledge that were previously unrelated to the paradigm of window-design should

be introduced. Here, the *function* is taken as a starting point with the *structure* being completely reimaged; both elements (subschemata) and their relationships (*structure*) are reconceived.

What is described here mirrors the effect of analogies. Goel and Bhatta (2004) argue that the use of analogous reasoning is the use of so-called *design patterns* which transfer content about *structure* (as defined in the FBS ontology) from source to target domain. *Design patterns* can thus be expressed in terms of schemata, where the design pattern is the selection of and relation between subschemata – the *structure* – mediating between *function* (dominating schema) and *behavior* (the dependent variable(s) in the dominating schema) (figure 27). Interestingly, analogies can thus be perceived as the knowledge component that allow humans to understand situations in domains that are beyond their existing knowledge, as analogies describe the content-independent logic behind the chain of causally related knowledge components (*structure*) that explains the link between object and behavior. Since the bottom-up approach is grounded in the notion of *paradigm-relatedness* (Nagasundaram and Bostrom, 1994-95) the quadrants visualized in figure 26 can be regarded as various types of *structure* compositions with varying levels of novelty. For example, ‘Redesign’ is the re-composition of existing subschemata, ‘Extend’ is the addition of new subschemata that were previously not part of the ‘paradigm’ of the domain of the specific design task, and ‘Transform’ is a combination of both recomposing and adding new subschemata. Boden’s types of creativity (2009) (figure 26) can be interpreted along the same lines. Analogies, being instances of structure that apply in both the source and design domain, can thus be regarded as vehicles for bottom-up design process augmentation.

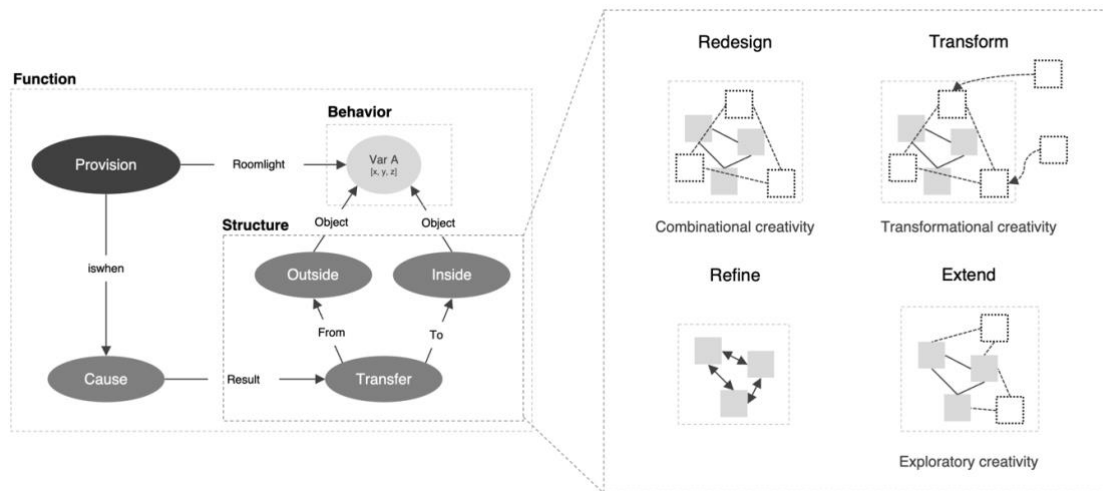


Figure 27: Schemata and types of novelty and creativity

### 6.5.2 Top-down novelty augmentation

A top-down version of novelty augmentation is one where the LLM generates novel ideas that inspire the designer and challenge it to step into new territory. Based on its

training data it would introduce examples from all sorts of different domains in which artifacts have been designed that somehow might be usable to the design question at hand. An example of top-down augmentation is the generation of the idea of a window that is capable adjusting its angle, and thus its volume of ventilation air, based on the required indoor temperature.

A more radical way to design a window based on knowledge that is far removed from the domains of the designer's knowledge could be to consider a design in which no glass is required, and hot and cold air are separated by internal flows of hot and cold air flowing through the walls of the house. This is a highly conceptual idea (it may not even work), but it may be very valuable to the designer in its thought-process, where after the interpretation it might trigger other domains of knowledge that would previously not have been associated with the design task.

As identified before, the value of the presented concepts depends on the interpretation of the designer based on his/her domain-knowledge. This presents a tension between the sorts of ideas and their corresponding domains (positioned either closely or remotely from the starting domain) that can be proposed as inspiration material and the domains of knowledge of a designer. A way to bypass the inherent limitation to the top-down approach, is to combine it with the bottom-up approach. Here, concepts grounded in unfamiliar (to the designer) knowledge domains are complemented by the associated 'knowledge elements', or ideas, and their relationships with both each other and the design question. In the case of the window concept, the augmentation system might use an analogy to explain the link between the idea proposition and the designer frame of reference. It can for example use the following analogy of a water dam: the window opens and closes just like a water dam might open and close its gates towards the hydro-electric generators based on electricity demand.

### **6.5.3 Directional novelty augmentation**

Directional novelty augmentation is to support the designer in its *evaluation* phase to reflect on the novelty of his proposition relative to his goals, as formulated in the activity layer of behavior (Cash & Kreye, 2017). This means, that the system proposes a reflection on the novelty of the designer's proposition(s) that affects the evaluation process of (explicitly or tacitly) (re)formulating goals, tasks, and actions. For example, given the design brief to design a window, where the design goal is 'to design the most elegant window at the lowest price possible' (see figure 20), and the executed design action is an idea proposition of such a window, the augmentation system may make suggestions about domains of knowledge relevant for the novelty enhancement of future propositions. An important note is that, although high levels of novelty may be seen as positive, there are also situations in which such radical choices are either not desirable or feasible. One may want to stick to convention or may not have the financial means to realize one-off radical designs. As such, there should be a way in which the designer can finetune the level of novelty.

## 6.6 Quality augmentation

Secondly, there is a direct connection between *usefulness* (or of a certain *quality*) and ‘of higher quality’. Following the definition presented by Dean et al. (2006), a quality idea should apply to the problem or design task at hand, is effective at solving it, and is implementable. A more detailed break-down of factors associated with quality is provided in figure 28. Here, quality is divided into *workability*, *relevance*, and *specificity* (Sääksjärvi & Gonçalves, 2018). As such, the quality of a concept can be enhanced when a designer is supported in crafting a design concept description that

1. clearly and exhaustively addresses the design brief on a functional level (*relevance*)
2. is shaped such that it complies with both social and practical affordances of the receiving party, *audience*, or material *context* (*workability*)
3. logically presents the connection between the features and capabilities of the product and the steps in which it causes the definition of status quo to be morphed into the definition of the desired state (*specificity*).

Each component is further subdivided in figure 28, but further defining each term is beyond the scope of this thesis. For now, it is relevant to observe that in order improve the quality of one’s ‘product’ it is critical to make sure that it is aligned with the initial design brief, the target audience, and that it is complete in its connection between the status quo and desired state. Following the structure for augmentation as presented in the novelty augmentation section, the content presented above is transformed in terms of the three different forms of quality augmentation.

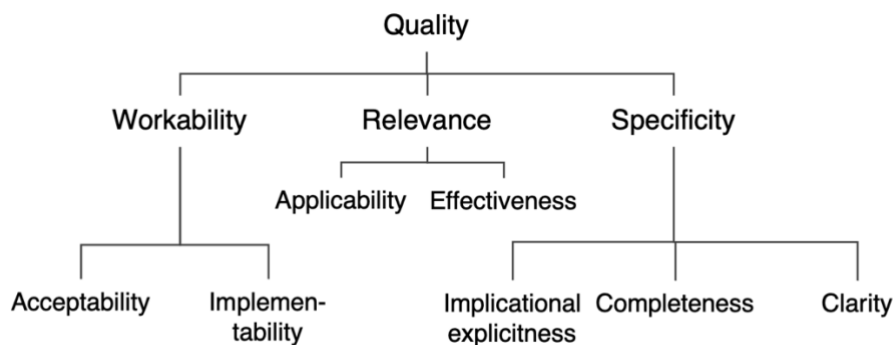


Figure 28: Definition of quality, adapted from Sääksjärvi & Gonçalves (2018)

### 6.6.1 Bottom-up quality augmentation

From a bottom-up perspective, quality augmentation implies that suggestions are made about knowledge elements to improve the connection between and completeness of causally related chains of knowledge components. This can be done by (implicitly) posing questions about certain knowledge domains or by probing the designer with various questions about completeness, logic, and contextual factors based on its understanding

of the design scenario, or knowledge components that together describe a logical chain of causes.

#### **6.6.2 Top-down quality augmentation**

A top-down approach to *quality* augmentation would generate a proposition for a design solution consisting of a logically structured and causally correctly proposed set of design contents that either have been produced by the designer or are part of the augmentation system. Similar to novelty augmentation, a form of hybrid quality augmentation may prevent any misunderstanding between the augmentation system and designer. Here, a quality concept is generated including the questions and considerations underlying the choices regarding its content and logic relative to the design brief.

#### **6.6.3 Directional quality augmentation**

A directional approach to *quality* augmentation implies that the designer is supported in its assessment of the status quo relative to the desired state such that the correct path emerges that connects the two. Directional quality augmentation is thus to support the designer in its *evaluation* phase to reflect on the effectiveness of his proposition relative to his goals, as formulated in the activity layer of behavior (Cash & Kreye, 2017). Here, the act of augmentation is focused on enabling the designer to spot misalignments between his previously set direction, as derived from his goals, and the interpretation of the design contents that result from it. The result can be the reformulation of goals, and consequent tasks, and/or the type of action that should be taken next. In addition, this type of augmentation also includes suggestions about domains of knowledge relevant for the quality enhancement of future propositions. In contrast to directional *novelty* augmentation, the focus lies on ensuring alignment with conventional wisdom to stay grounded in reality and make sure that other stakeholders are represented in the final solution. The combination of both directional *novelty* and *quality* augmentation forms a nice complementary mix of free exploitation of all possible options, and process-related practicality and trust, to ensure the most creative outcome and as such a successfully augmented process.

#### **Takeaway box 7**

Design process augmentation that focuses on the enhancement of novelty comes in three increasingly novel variants: combinational, exploratory, and transformational. These variants use analogous reasoning, as captured in design patterns, to extend the individual's ability to recall and use both novel knowledge components and to construct novel relations between these components. In the bottom-up form, this means introducing these components and relations. In a top-down scenario, these mechanisms are used by the augmentation system to generate novel propositions that inspire the designer. In directional augmentation, the system proposes knowledge components (with varying degrees of novelty) to influence the potential

novelty of the design process resulting from the (re)formulation of goals, tasks, or actions.

With regards to quality augmentation according to the bottom-up type, suggestions are made about knowledge elements that improve the connection between and completeness of causally related chains of knowledge components. A top-down variant proposes a logically ordered and complete chain of knowledge components from the status quo to the desired state. In *directional* quality augmentation, the designer is supported in its assessment of the status quo relative to the desired state such that the correct path emerges that connects the two.

Building on the process-related creativity literature, each of the six forms of design process augmentation can be described as a set of process characteristics. A brief overview of the various ways in which novelty and quality augmentation of the design process can be established along the three types of design process augmentation is provided in table 11.

Engelbart intellect augmentation	Enabling what was not possible before	Higher quality
Creative product	Novelty	Quality
Bottom-up (Interpretation)	Introducing 1) new (to the designer) knowledge elements and 2) presenting new ways in which these knowledge components can be related.	Introducing missing components in the chain of logic or probing the designer with various questions about completeness, logic, and contextual factors.
Top-down (Proposition)	Generating novel (to the designer) concepts that inspire and challenge the designer's interpretation.	Generating a logically structured design solution description based on all the design contents generated by the designer and LLM
Directional (Goals & evaluation)	Proposing a reflection on the novelty of the designer's proposition(s) that affects his evaluation process of (explicitly or tacitly) (re)formulating his goals and tasks and providing novelty-enhancing suggestions for further action.	Enabling the designer to spot misalignments between his previously set direction, as derived from his goals, and the interpretation of the design contents, to better inform the definition of his goals, and providing quality-enhancing suggestions for further action.

Table 11: Summary of design process augmentation

An important consideration is that the success of the augmentation during the process as defined by the outcome of that process should be considered a continuum; the novelty of a generated concept proposition, or knowledge component can vary.

Moreover, for the proposition to augment a specific designer, it depends on the designer's knowledge whether the novelty and quality of the proposition can be considered 'augmenting' given a certain design context. As such, it is important that the augmentation system knows the designer's relative strengths, his knowledge domains and expertise, and weaknesses to adjust the way in which it interacts.

### 6.7 Embedding the augmentation forms in a collaborative process

The previous subsection has established an overview of the six forms of design process augmentation as defined at the scale of a mental iteration. This section aims to translate these process-related definitions of augmentation to a formulation of a human-AS (Augmentation System) collaborative design process framework. The central question is how these forms of augmentation when executed by an Augmentation System (AS) can be usefully implemented in the design process. For example, how and when does the Augmentation System (AS) respond to a designer who proposes a certain idea? This question is especially critical since the type of response of the AS depends on 1) the mental iteration stage that the designer aims to augment (*interpretation* = top-down, *proposition* = bottom-up, *evaluation* = directional), and 2) the focus of the augmentation on either *novelty* or *quality*. Moreover, as mentioned before, design actions can be viewed from the perspective of both the human and CA. As such there is third category that increases the number of variables that determines the type of AS response: the type of design action (*representation*, *information*, *knowledge sharing*) of the AS that is desired given the scenario. As observed in section 5.11.2, *representation* action is part of every action type. As such, the three types of AS-based design actions, as listed in table 12, are packaged into two types resulting in twelve different ways (three augmentation types \* two types of augmentation focus \* two types of design actions) in which the augmentation system can respond. The combination of augmentation form and design action is called an *augmentation move*. This term also aligns nicely with the language used in the Supermind Ideator, another design augmentation system based on LLMs, where the term *design moves* is used to refer to the actions of the system (Rick et al., 2023).

Representation	Information	Knowledge sharing
The augmentation system generates an (alternative) representation of the design process related content	The augmentation system provides an alternative composition of the data and additional data.	The augmentation system explains, through reasoning, its perspective on the input representation based on its understanding and

Table 12: The design actions as medium for expressing the types of augmentation

The question then becomes how the human and AS are supposed to derive the internal state of each other (this refers to the actor-related category of 'perception' from the framework of Viros-i-Martin and Selva (2021) (table 3), which informs the type of



response that is required, or in the case of the AS its *augmentation move*. To enable the AS to infer the internal state of the designer, it can be said that the designer should always inform the AS about his goal (*activity layer*), or general objective, his task (*task layer*), and the type of design action that is performed (*action layer*). Similarly, the AS should also provide such context when responding to the designer. However, in practice, this interaction is too cumbersome. Besides, the AS still does not have enough contextual information to infer what form of augmentation or type of design action is most useful at a certain point.

Instead, it would be more elegant to work towards the notion of participatory sense-making from (Davis et al., 2015). They argue that participatory sense-making is useful to analyze, understand and model creative collaboration. In the words of De Jaegher and Di Paolo (2007, p. 493): [it occurs where] "*a co-regulated coupling exists between at least two autonomous agents where the regulation itself is aimed at the aspects of the coupling itself so that the domain of relational dynamics constitutes an emergent autonomous organization without destroying the autonomy of the agents involved*". To work towards this scenario, two interventions are created. First *personas* are presented as a way to embed the contextual information regarding the desired form of augmentation. Second, different *interaction views* are introduced within which the personas can reside. Combined, the *personas* represent the actual act of augmentation, whereas the *interaction views* are the different media via which these acts of augmentation are materialized. Both interventions are discussed in the following two subsections.

#### 6.7.1 Personas as vehicles for the communication about augmentation forms

A persona is defined here as a collection of static characteristics grounded in the six forms of augmentation. As such, personas make the process of inferring the other actor's internal state obsolete. By attending to a specific persona, the designer knows the internal state of the system in the instance of that specific interaction, as he himself has composed the persona and thus the 'internal state'. For example, in the framework of human-AI collaborative design space exploration, the internal state of the AI-based system, or CA, is described by its role. Here, examples include the role of a historian, analyst, explorer, expert, or critic. The various 'roles' as part of the internal state of a CA (table 3) can be expressed in terms of a combination of augmentation type and focus. From now on, these roles will be referred to as *augmentation personas*. The 'analyst' *augmentation persona* can be configured as visualized in figure 28.

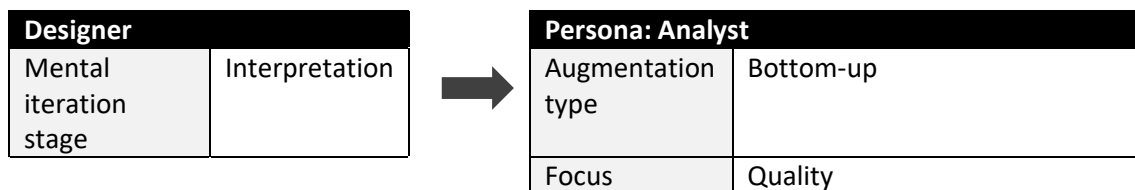


Figure 28: Analyst role expressed in a persona

The implication is that instead of having to select or mention the mental iteration one would like the AS to augment, one can set the (temporary) internal state of the AS by configuring roles or personas, based on the six forms of augmentation specifically tailored to each mental iteration stage. For example, a designer might configure the analyst persona (figure 28) when he wants the interpretation stage to augmented (bottom-up) in a way that improves the internal logic (quality) by having the persona generate an alternative representation of the information and explain its reasoning (*representation + knowledge sharing* action; see interaction views for a more detailed description of this part). By moving beyond the notion of ‘role’ to that of ‘persona’, characteristics from the buckets of ‘attributes’ and ‘knowledge’ can also be considered part of the configuration. This insight is linked to the framework of Viros-i-Martin and Selva (2021) in table 13.

Human-AI collaborative design space exploration framework			Mental iteration design process & LLM capabilities
Inferred state of the CA/human	Human	The expected interpreted state of the CA given the interaction	The choice for a persona determines the state of the CA
	CA	[Not defined, implied to be similar to the human]	The choice for a persona determines the state of the human

Table 13: Personas related to the collaborative human-AI design space exploration framework (Viros-i-Martin & Selva, 2021)

### 6.7.2 Two interaction views

To ensure the natural integration of the three types of design actions, as packaged in combinations of *representation + information* action or *representation + knowledge sharing* action, via which the AS can execute the six forms of design augmentation, two types of *interaction views* are introduced. As identified in section 5.11.2, both *information* and *knowledge sharing* action are always accompanied by *representation* action. Therefore, the three types of design actions are translated into two *interaction views*. Just like with *augmentation personas*, the choice for either one of these views makes the process of inferring the other actor’s internal state obsolete. These *interactions views* represent different ways in which one can consume the representation actions of the AS, since representation action is the common denominator in both views.

#### *Information-organization interaction view*

First, *information* action is embedded in an *interaction view* that presents the content that is being worked on by the user and generated through a collaborative process with the *augmentation personas*. Besides, this *interaction view* offers the option to organize the design content in various compositions. Here, the interface may restructure the

design content such that new, and previously unidentified, dimensions become visible. This *information* action is carried out through a form of *representation* action, projecting the data parts onto a certain medium.

#### *Perspective interaction view*

Second, *knowledge sharing* action is embedded in an *interaction view* that presents a certain perspective on the set of design process components. In this view, the AS may provide suggestions to enhance either the quality or the novelty of the knowledge domains and/or ideas in the next sequence of actions. Again, this type of action is accompanied by *representation* action, since the suggestions are again projected onto a certain medium.

#### **Takeaway box 8**

To translate the six forms of design process augmentation to a sequential form of collaboration, two interventions have been introduced that facilitate the interaction between human and AS. First, *personas* are an embedding of the six types of design process augmentation. Second, the two *interaction views* are an embedded way of communicating about the desired form of design action an actor expects as a response. Both interventions describe how the human and AS infer each other's internal state, leading to an effective collaboration.

### **6.8 Conceptual framework for human-AS co-design**

To consolidate the findings of this thesis, a framework (figure 28) has been constructed that represents this logic and combines it with the previously discussed layers of design activity and the mental iteration design process framework to create a time-based conceptual framework of a GenAI-augmented design process, where an LLM is a specific type of generative AI that is considered feasible to use for the twelve augmentation moves. The visual also makes explicit how collaboration in general is such a powerful way of working, as one doesn't only interpret one's own output, but also one's output transformed by another actor that has applied certain transformations to it such that extra material is presented that can help to further progress one's next mental iteration. The framework presented in figure 28 can be regarded as the fundament for further conceptual expansion of current body of design theory towards the collaborative end of the design-spectrum, as such meeting the demand for conceptual expansion towards the collaborative domain identified by (Crilly, 2021b) and (Cash et al., 2023). Moreover, it forms the basis for the interface proposed in section eight. Figure 29 shows a hypothetical progression of design process in a designer-augmentation persona interaction. Here, the *augmentation persona* is configured as an 'analyst'.

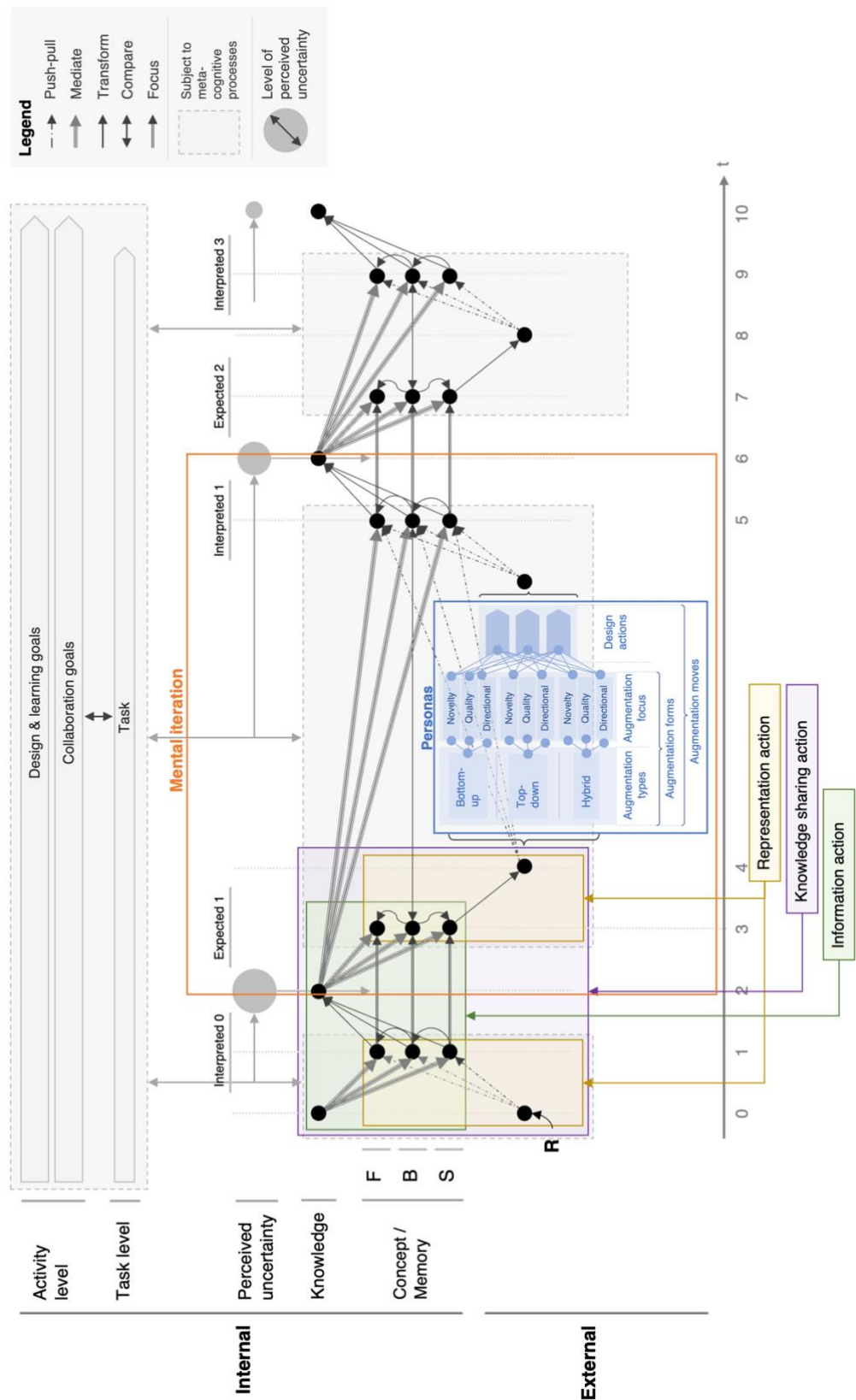


Figure 28: Conceptual framework of human-AS augmented co-design

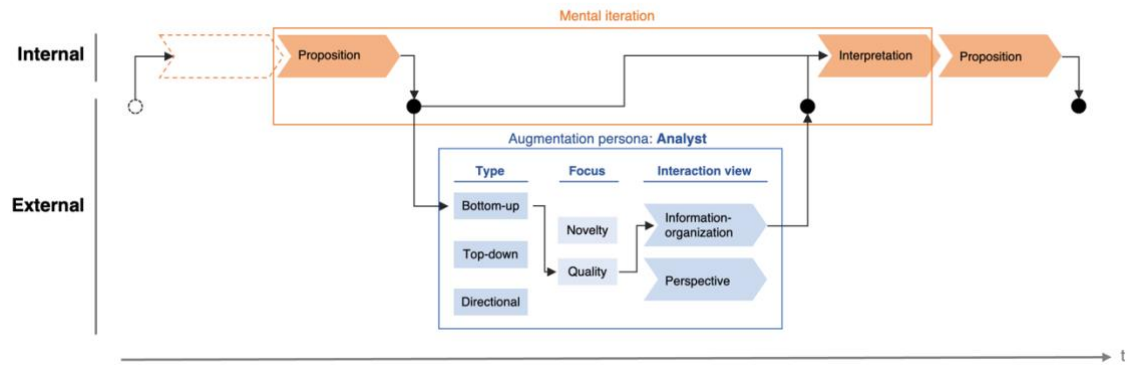


Figure 29: Design process progression with an augmentation persona

## 7. Design augmentation capabilities of LLMs

---

Language has been one of the earliest forms of intellect augmentation and has enabled humans to more expressively and at increasingly high levels of resolution communicate. Later, script increased the scope of an individual's memory to that of many later generations. With regards to script, a combination of certain methods and artifacts enabled humans to capture language in a certain medium. Throughout time, these media have been transformed to today's digital interfaces, such as the internet that completely changed the way in which written accounts of language could be distributed. Given the widespread distribution of information via the internet, the internet has accumulated vast amounts of human-created accounts of information. This data now represents the source of data that powers today's LLMs. The question arises if the characteristics of this information and the way in which it is embedded into LLMs is a good fit with the design process. This leads to the third research question:

*What are the capabilities of LLMs to realize design process augmentation practices?*

This question is relevant as it provides the theoretical foundation for describing the Augmentation System's (AS') capabilities when LLMs are used as the technological fundament. The insights can be used to detail the AS' (or CA's) characteristics as defined in the collaborative human-AI design space exploration framework (Viros-i-Martin & Selva, 2021), just like section 5.16 provided this definition for the human actor in the process of design. This section thus forms the last piece of the puzzle in conceptually describing the construct of LLM-based design process augmentation.

---

### 7.1 Approach

As identified in the previous section, the qualities associated with successful augmented outcomes – being of 'higher quality' and 'enabling what would not be possible before' – reflect the qualities of what is described as creative product – a product is creative when it is both novel and useful or is of high quality. Getting to such a product in a way in which the designer is augmented by an LLM means that the LLM is required to fuel the designer with information and ideas that themselves are novelty and quality enhancing

relative to the designer's capacities. As such, the LLM is expected to be creative itself, since it has to interpret, manipulate and thus creatively process data parts, both from itself as well as from the designer. Therefore, an account of what LLMs are capable of relative to the creative process in design must be established, as the augmentation can be regarded as a creative process in itself. Attending to the various augmentation approaches, five questions arise:

#### *Novelty*

1. Bottom-up: Are LLMs capable of introducing knowledge components and propositions to increase the novelty of the designer's output?
2. Top-down: Are LLMs capable of producing novel concept propositions relative to the context in which it is externalized?

#### *Quality*

3. Bottom-up: Are LLMs able to evaluate the workability, relevancy and specificity of the total of design ideas, referred to as the quality of the solution?
4. Top-down: Are LLMs capable of producing a quality (workable, relevant and specific) and logically structure proposition based on its own knowledge and content shared by the designer?

#### *Directional*

5. Novelty & Quality: Can LLMs coherently progress through an iterative process in which it enables the designer to arrive at new insights?

The following subsections aim to answer these five questions. First a conceptual description of the inner mechanisms of an LLM is provided. Second, building on this basic understanding of how LLMs function, a set of hypotheses on the design process augmentation capabilities of LLMs is generated by reflecting on the creativity of LLMs from the four perspectives on creativity: *person, press, product, process* (Rhodes, 1961). This forms the basis for the further specification of the AI-related categories in the human-AI design space exploration framework (Viros-i-Martin & Selva, 2021), and as such completing its translation to the domain of this thesis: LLM-based design process augmentation. Finally, two practical examples are proposed to partially reflect on the validity of these hypotheses.

## **7.2 A conceptual understanding of how Large Language Models (LLMs) work**

To understand what LLMs are and how they work, it is relevant to understand their position in the larger context of Artificial Intelligence (AI). AI can be described as "*an area of computer science that involves building machines capable of performing tasks which require human intelligence*" (Kalyan, 2024, p. 2). AI comprises the fields of Machine Learning (ML) and within it Deep Learning (DL) (Kalyan, 2024). Positioned within the latter, LLMs are a special class of Pretrained Language Models (PLMs) based on a Transformer architecture. Figure 30, adapted from Kalyan (2024), visualizes the position of LLMs relative to the wider field of AI.

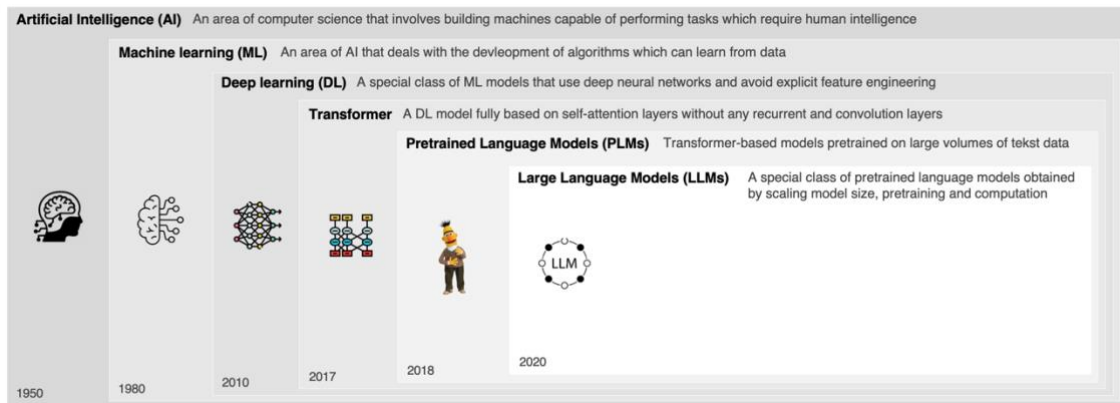


Figure 30: Positioning LLMs relative to the broader field of AI (adapted from Kalyan, 2024)

At a high level, LLMs generate language that is predicted to be most likely to follow up on the input text, called the prompt. Three stages can be identified that describe how this process works: 1) tokenization, 2) encoding, and 3) decoding. Each stage will be detailed below. Given the purpose of this thesis, it is relevant to understand how LLMs generate language at a conceptual level. The visualizations that complement the textual explanations should therefore be interpreted as illustrative and simplified representations of the actual mechanism of language generation in LLMs.

### 7.1.1 Tokenization & pre-use pre-training

Each input prompt, a collection of sentences and words, is dissected into *tokens*. *Tokens* are words or parts of words that become the basis for the construction of vector representations that aim to capture the ‘meaning’ of these tokens based on their connection relative to the rest of tokens within the dataset (Kalyan, 2024). To illustrate this, figure 31 visualizes tokens, chosen at the scale of words or other linguistic components.

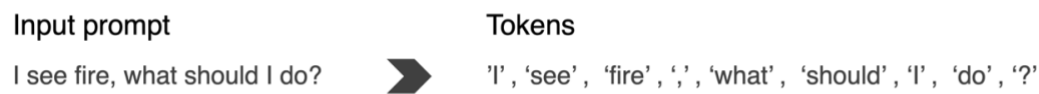


Figure 31: Tokenization

The vector representations of tokens are constructed during the first phase of training, called pre-training. During pre-training, the transformer is trained on unlabeled textual data, following the self-supervised learning paradigm (Kalyan, 2024). The result is a model that has constructed so-called ‘embeddings’ of each token that capture its semantic and syntactic features (figure 32); the transformer, now called a Pretrained Language Model (PLM), has created a vector-based representation of its training dataset.



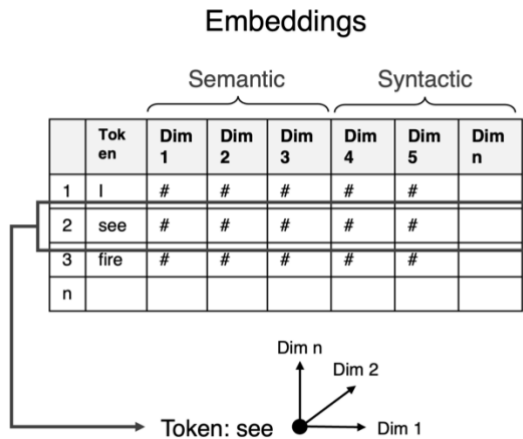


Figure 32: Token embeddings

### 7.1.2 Encoding

During encoding, the set of tokens is transformed into rich contextualized vector representations for each token, that capture both semantic and syntactic features of the token (Kalyan, 2024). This vector representation is constructed by feeding the prompt forward through a set of encoder layers. Each encoder layer includes an ‘embedding layer’ and a ‘self-attention mechanism’. The embedding layer transforms input tokens into the corresponding vector representations (figure 32). The ‘self-attention mechanism’ ‘adds contextual information to each token’s vector representation by capturing the relative strength of the relation between the token’s in the input prompt based on their vector representations (figure 33) (Kalyan, 2024; OpenAI et al., 2024).

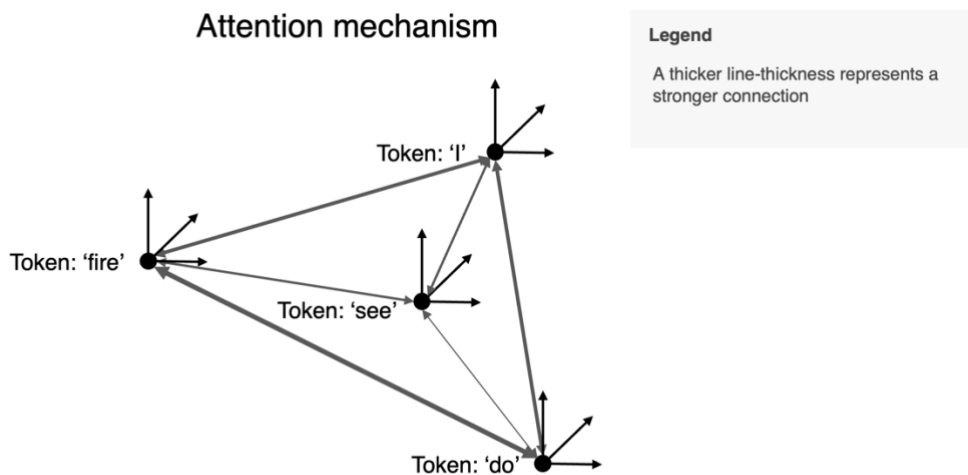


Figure 33: Encoding based on word embeddings and the attention mechanism

### 7.1.3 Decoding

Based on the encoded tokens from the input prompt, the decoder – a stack consisting of an embedding layer, decoder layers, and an output layer – generates the response to the input prompt. Each decoder layer comprises two attention mechanisms (figure 34). First, ‘encoder-decoder attention’ ensures that the output attends to the original input prompt (Kalyan, 2024). Second, the ‘masked self-attention’ verifies whether each generated token attends to the previously generated token to make sure that the output is a semantically and syntactically correct linguistic proposition (Kalyan, 2024).

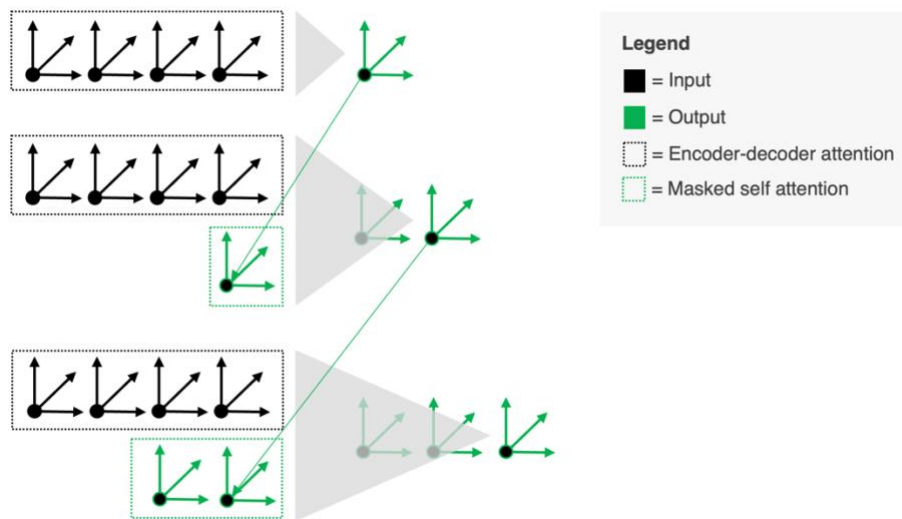


Figure 34: Decoder layer attention mechanisms

Based on these attention mechanisms in combination with the token embeddings, the output layer predicts what token, or set of tokens, is most probable to follow up on the input prompt. One can conceptualize this process as chance-based mapping between vectorized input tokens and all other vector-embeddings of the LLM, as visualized in figure 35. Figure 35 shows the chance-based mapping of tokens by looking at the token ‘fire’ from the input prompt. This process is called next-word prediction. The output may start with the token ‘exit’ followed by ‘the house’.

#### Takeaway box 9

LLMs are part of the class of generative AI (GenAI) models. They generate language by 1) dissecting the words from the input prompt into tokens, 2) translating symbol-based tokens into numeric vector embeddings, 3) calculating the relational strength between the vector embeddings to identify the relative relevance of a certain token in the sequences of tokens, and 4) predicting the most probable set of next tokens based on both the input tokens as well as the already generated response tokens.

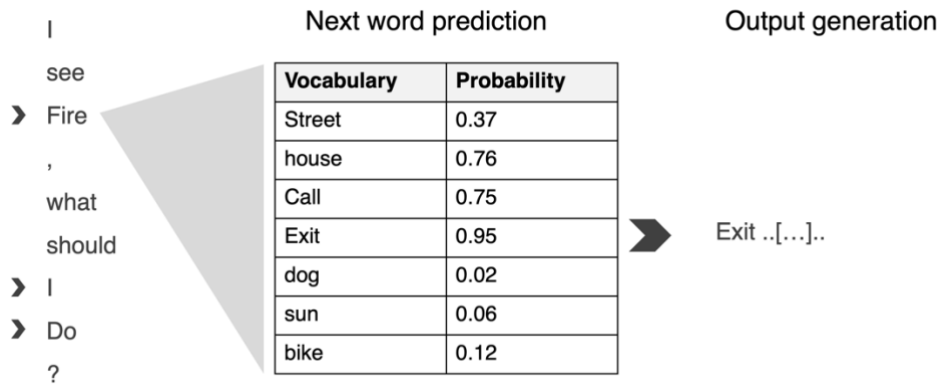


Figure 35: Next-word prediction

### 7.3 Theoretical approach: LLM augmentation novelty, quality, and iteration

Building on the conception of the mechanisms underlying LLMs as described above, this section aims to contribute to answering the five questions about the LLMs ability to augment the design process from a creativity theory perspective. Recalling from section six, this domain of literature has been chosen for two reasons. First, the definition of a creative product in creativity literature resembles the definition of human intellect augmentation from Engelbart (1961). Given the three other perspectives represented in creativity literature (*person, press, process*), Engelbart's output-oriented definition of human-intellect augmentation can be translated to process-, person- (or computational actor), and press-related attributes. This provides a more holistic and complete view on the capabilities of LLMs in the domain of design process augmentation. Second, design can be researched through the lens of creativity and vice versa (Taura & Nagai, 2010). Creativity is the word used to describe the underlying capabilities that enable a process of design, and as such form a good starting point for assessing the design process augmentation capabilities of LLMs.

Current empirical research that has assessed LLMs on their ability to reason (Hao et al., 2023; Qin et al., 2024; Wang et al., 2024) or to perform creativity assessments such as the Alternative Uses Test (AUT) (Stevenson et al., 2022) concludes that LLMs can perform or simulate these mechanisms, although not yet at the human level. Others have taken a more holistic and theoretical approach to evaluate and challenge existing theories about human creativity by comparing the nature and inner mechanism of AI systems to the reigning conceptions about the facets involved in human creativity (Boden, 1998, 2009; Runco, 2023b; Runco & Jaeger, 2012). Their findings pointed at the distinction between artificial and human creativity being the presence of intrinsic motivation (Runco, 2023a), something currently lacking in models such as LLMs among many other AI systems. More specifically applied to LLMs, Franceschelli and Musolesi (2023) argue that LLMs are not creative at the transformational level (Boden, 2009) as they cannot update their fundamental knowledge structure. However, a holistic theoretical review of LLM creativity that approaches LLMs from all perspectives associated with creativity – such as the four Ps (person, product, process, press) of

creativity (Rhodes, 1961) – still lacks, as such hindering our ability to answer the three question stated at the beginning of this section. For example, current investigations into LLM creativity (such as the ones mentioned above) focus mainly on the *process* or *product* related creativity characteristics of LLMs (Jiang et al., 2024) without a clear account of the LLM as being a creator given a certain context. The *person* and *press* perspectives of the four Ps of creativity framework (Rhodes, 1961) have mostly been neglected.

As such, the following subsections start by interpreting LLMs from the perspectives of *person* (or *actor* (5As), *collective* (5Cs)) and *press* (or *audience & affordances* (5As), *context & community* (5Cs)), as this results in a more intuitive conception of how one can look at an LLM as being a perceiver and producer of content, embedded in a certain context. Secondly, building on this conception of the LLM as a creative (to a certain extend) actor, the *product* and *process* perspectives, both highly intertwined, bring forward the initial hypotheses on the LLM's ability to generate novel, and quality domains of knowledge and ideas in an iterative way.

### **7.2.1 The conception of an LLM as an actor in an environment**

From the perspective that is concerned with the characteristics of the entity that generates the content, aspects such as personality traits, attitudes, intelligence, habits, as well as behavior such as curiosity, independence, persistence, and openness among others are identified as the characteristics of creative (read 'novel') individuals (Stein, 1968). A characteristic that is related to curiosity is 'intrinsic motivation' (Rhodes, 1961). Closely related to motivation is intention, something that according to Runco (2023a) is not present in any current form of AI as it implies intrinsic motivation, and authenticity. The latter two factors have been added to the 'standard definition' of creativity, to make a distinction between human and artificial creativity (Runco, 2023a). Similarly, according to Franceschelli and Musolesi (2023) LLMs do not contain intrinsic motivation as it requires an interest in a task (Deci & Ryan, 1985). They merely generate a string of tokens resulting from the input set of tokens.

In the context of Computational Creativity (CC), Jordanous uses the word *producer* instead of *person* or *actor* (2016). The computer, or more accurately the computer algorithm, software, or creative agent, is considered to be the producer, where the computer hardware and software can be compared to the physical and functional characteristics respectively of a person (Jordanous, 2016). The algorithm is embedded in the physique condition of the computer. An additional nuance to the definition of the *producer* is the notion of the programmer(s), tester(s), researcher(s) and the individual(s) that interact with the computer system being included in what is described as the *producer* (Jordanous, 2016). This implies that the computer program is merely an extension of the human creativity and not a creative entity itself.

This view is also reflected by the 5C framework in which the *person* is replaced by the *collective* of both humans and computers (Kantosalo & Takala, 2020). This puts more emphasis on the user as bringing and/or implementing a complementary set of creative attributes to the entity – the LLM for example – that is being assessed on its creativity. This notion that the external context effectively shapes the individual is also manifested in the 5A framework of creativity, where Glăveanu describes the *actor* as “*being defined by a system of social relations and cultural traditions regulating these relations*” (2013). The actor is not only an individual with a set of cognitive functions and mechanisms, but a reflection and embodiment of social and cultural conventions learned and abstracted from its environment. Similarly, Jordanous (2016) argues that the computer algorithm, or LLM, is the embodiment of its creators. By redefining *person* into *actor*, Glăveanu aims to strike a balance between the inherent bias of a framework towards the exclusive focus on only one perspective and therefore the sole attention on the individual. As mentioned by Glăveanu (2013) “creativity relies on the individual”, but “individuals are also ineluctably social and cultural phenomena” (Markus & Hamedani, 2007, p. 5).

Building on the embedded concept of the actor, an LLM can be considered a conversational representation of human knowledge and culture extracted from textual training data and expressed through natural language. This data, created by humans, is the human experience of its environment captured in natural language. Building on the conventions of ‘cultural psychology’, one could consider the LLM’s fundament – the generative pre-trained transformer – to *be* the embedding of all social and cultural language-based contexts. This closely corresponds to the terminology used in computer-science for the representation of tokens, (parts of) words, as ‘vector- or word-embeddings’ (Kalyan, 2024). Here, word embeddings are multidimensional vector spaces representing language, allowing vector operations to express semantic relationships between words (Mikolov et al., 2013).

From the perspective of *person, actor or collective*, creativity (of which novelty is a large part) demands intention, intrinsic motivation and authenticity (Runco, 2023b) where motivation is considered to be related to specific behaviors such as curiosity and persistence (Rhodes, 1961). LLMs do not possess such characteristics, if compared against their fundamental ‘modus operandi’ of next-word prediction. However, there are, ways in which such characteristics can be simulated through prompting. Although this can be considered part of the *process* perspective, it will be investigated in this section as it is a way to alter the characteristics connected to the *person* perspective.

First, the Autonomous Cognitive Entities (ACE) framework proposes a way to simulate motivation, using the LLM’s language interface (Shapiro et al., 2023). The framework, presented in figure 36, consists of six layers that form a constant cycle of top-down control and bottom-up data. It introduces the ‘aspiration layer’ that includes a *mission, values, morals* and *purpose* – the latter one being a factor that according to Rhodes (1961) is highly correlated with the factors that are said to influence the creative

process. The section about *process* will further continue this analysis by highlighting the steps in the creative process that follow after motivation: response generation, validation and communication (Amabile, 1983). In short, this framework uses the LLM as an interface to simulate aspects of motivation through a mission and purpose, but still lacks authenticity (Runco, 2023b) as those factors are not learned by the model itself but provided by a human.

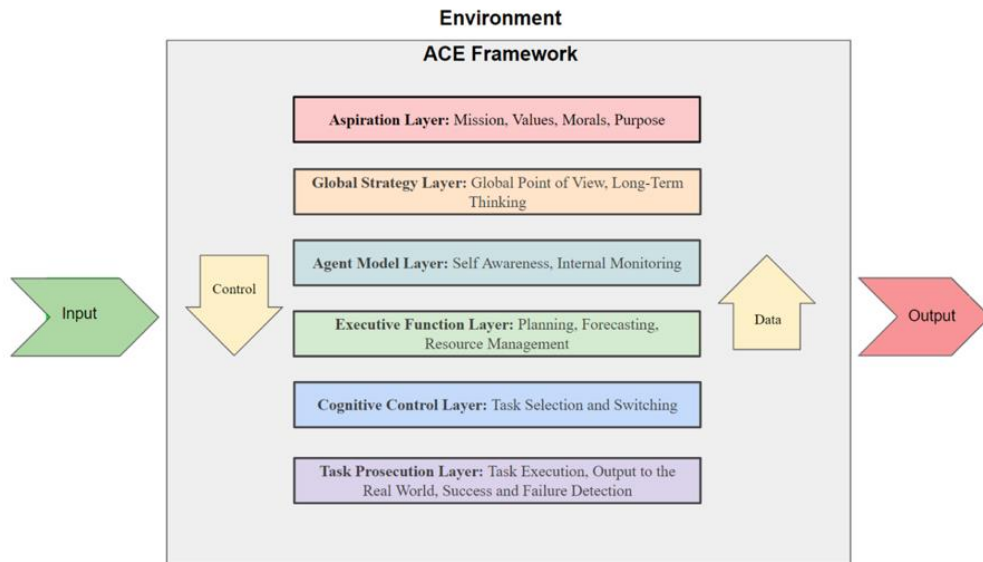


Figure 36: Autonomous Cognitive Agent (ACE) framework (Shapiro et al., 2023)

Second, user-prompting is a more feasible way for users to alter an LLM's characteristics. Recent research has identified that prompting an LLM to exhibit certain personality traits – this study tested neuroticism, extraversion and agreeableness – can alter the opinion of an LLM on given questions (Mao et al., 2023). Although the LLM does not have personality traits, as defined in the psychological field, it can simulate these by considering the tokens that represent the personality trait – in the form of word-embeddings – in its prediction of next output-tokens.

Finally, research efforts are currently exploring pathways to use finetuning to adapt LLMs to specific use cases and tasks (Franceschelli & Musolesi, 2023). One could imagine finetuning to be used to alter the current characteristics of the LLM and to optimize it for creative performance. Here, one LLM could be finetuned and prompted to generate outputs that are a little less concerned with the *quality* (defined in this thesis as a combination of *workability*, *relevancy*, and *specificity*) and focus mostly on the *novelty*, while afterwards another LLM that is finetuned for *quality* can be used to ensure the utility and value of the idea within a certain situation. Although this might allow the LLM to explore less-probable tokens, the current LLMs will not be capable of transforming their fundamental rules underlying the exploration in generating language (Franceschelli & Musolesi, 2023). To be considered creative, or novel, at a

transformational level (Boden, 2009), an LLM requires the capability to iteratively and constantly re-finetune itself based on new interactions and inputs. Techniques such as continual learning (Kirkpatrick et al., 2017; Shin et al., 2017) for LLMs (Sun et al., 2019; Wu et al., 2022) are potential, yet unexplored ways to realize this (Franceschelli & Musolesi, 2023).

#### Takeaway box 10

LLMs do not possess the characteristics that are related to creative individuals, such as (among others) curiosity, persistence (Rhodes, 1961) and motivation (Runco, 2023). Although proposed in a tentative way, it could be argued that both the factors *novelty* and *quality* may be affected by the lack of these characteristics, as higher levels of curiosity may lead to increased *novelty* and more persistence to higher *quality*, fueled by the presence of motivation. A well-orchestrated interaction and workflow between the LLM and the designer, however, might compensate for this. As will be touched upon more in the section about the *process*, the personal characteristics related to creativity can be simulated to a certain extent via a hierarchically layered operating model (ACE-framework), user-prompting, and finetuning. However, since the LLM has not autonomously learned the ‘reasoning’ or ‘thinking’ patterns presented in these prompts and frameworks, its creativity cannot be considered ‘authentic’<sup>11</sup>. But, since the creativity of LLMs is not the focus of this section per se, instead using the creativity theory to assess the LLM’s capabilities in the domain of augmentation, that is not critical.

#### 7.2.2 Defining the context that has shaped the LLM during training and affects it during use

Given our view on LLMs as being the embedding of their training data, it is relevant to look at the context that has surrounded LLMs during training and that surrounds it after training. The *press* perspective allows us to do so, as it comprises the physical and social environment in which the *actor* – in this case the LLM – is situated (Rhodes, 1961). This section will first establish a general overview of creativity literature from the *press* perspective to then interpret it in the light of LLMs.

The *press* houses the actor and the actor’s externalized or embodied outings – sentences or formulations of ideas can be considered an LLM’s embodied outings – with the coordinated set of interactions both within the actor as well as in between the actor, its *products* (next section), and the environment representing the *process* (Rhodes, 1961). All components can be regarded as inter-related and co-evolving together, through reciprocal influence (Glăveanu, 2013; Halstrøm & Galle, 2015; Hatchuel & Weil, 2002). The environment – or the *press*, *audience*, and *affordances* (5As), or *community*

---

<sup>11</sup> The introduction of OpenAI’s o1 model may change this, as it has trained itself on strings of reasoning steps generated by itself.

*and context* (5Cs) – exerts influence on the (set of) *person(s)* or *actor(s)* within it, manipulating the attitude and characteristics of those and thus the shape and content of their *product*. The actor's *action* then results in a *product* that expands the environment while getting feedback in the process of doing so (Glăveanu, 2013).

Glăveanu (2013) has divided the *press* perspective into *audience* and *affordances* highlighting the importance of a divide between the social and material environment. Glăveanu argues that *audience* is preferred over *press* as it brings to life the persons that might assist, judge, or even contribute to the creative and cultured *artifacts* (2013). By positioning both active and passive participants within the *audience*, the act of creating and perceiving are considered equivalent. Similarly, Dewey (1934) states that “*to perceive, a beholder must create his own experience. And his creation must include relations comparable to those which the original producer underwent*” (p. 56). Continuing this line of reasoning, Dewey notes that during the creative act, “*the artist has to become vicariously the receiving audience*” (p. 111). The audience, whether in the physical sense or as mental representation, therefore guides the dialogical process of creativity (Grossen, 2008). With *community*, the 5C framework (Kantosalo & Takala, 2020) only refers to the passive audience, citing evidence that points at the different attitudes of humans resulting from either a role within the audience or actively collaborating collective of actors (Bown, 2015).

Glăveanu's *affordances* focus on the structure that the physical objects in our world provide to our *actions*, referring to ‘culture’ as something that mediates our interpretation of the potential affordances that the physical objects could provide us (Glăveanu, 2013). To exploit the potential that resides inside the physical environment – or *affordances* – the *actor* requires certain knowledge and abilities (Glăveanu, 2013). Interesting to the investigation into the LLMs' ability to explore such novel potential is the question where such knowledge originates from. However, Glăveanu (2013) doesn't state where this knowledge originates from, apart from the notions that it can be used to “*discover new affordances, and even “create” the ones needed to fulfill a specific action*” (p. 76) and that “*individuals [...] need a considerable amount of time to learn and practice the actions expected of them before making clear contributions to the knowledge and practices of the group*” (p. 72).

The 5C framework addresses this issue by describing the context as “the materials with which the collective interacts during co-creation, the previous influential works the collective may draw inspiration from, and the cultural norms and rules which may affect the collective and its work” (Kantosalo & Takala, 2020, p. 21). The second element in this enumeration refers to the concept of knowledge, which computational creativity is called the inspiring set (Ritchie, 2007). The inspiring set indicates the subset of items that together construct the computer program that is assessed on its creativity, and includes items that the program is designed to replicate or knowledge bases which drive the computation within the program (Ritchie, 2007).



It is important to note that both authors only consider the effect of the *press* on the creativity of the *actor* during the creative performance, with the only exception being the inclusion of the *inspiring set* in the definition of *context* in the 5C framework (Kantosalo & Takala, 2020). It is therefore crucial to make a distinction between the effect of the environment on the LLM during training and the LLM during the use phase, as LLMs have a data cut-off point that leaves its knowledge of the world locked in a particular state (OpenAI et al., 2024).

First, in defining the effect of the *audience* or *community* on the LLM during the training phase, it could be said that the LLM becomes or embodies its *audience* and *affordances*, or training data while ‘interacting’ with it. The LLM extracts the relevant parameters from its training dataset – the *inspiring set* (Ritchie, 2007) – and its interactions with human feedback (during Reinforcement Learning from Human Feedback (RLHF)) while constructing the word embeddings that allow it to generate natural language and that represent a derivative of its *inspiring set*. The human evaluators who train the reward-model that finetunes the LLM during RLHF, can then be considered active contributors to the LLM’s ability to generate natural language that is human-aligned. After being trained, the LLM has embodied the values and rules extracted from the *audience*, the human evaluators during RLHF, and is then capable of operating in line with Dewey’s description of a creative *actor*, who has become one with the *audience* in the process of conceiving the creative object (1934). This also follows Glăveanu’s view of the *actor* as the embodiment of social and cultural conventions (2013). An example of this ability to empathize with the audience is provided in table 14. The LLM can be instructed to generate language that is situated within the space of knowledge and experience of a specifically defined audience, similar to prompting the model to display certain personality characteristics (Mao et al., 2023).

Audience	Plumber	7-year old
Prompt	Please explain how a toilet works to a plumber. Formulate your answer using at most 70 words.	Please explain how a toilet works to a 7-year-old. Formulate your answer using at most 70 words.
Response	A toilet operates on the principle of a siphon, initiated by a flush that releases water from the tank into the bowl via the flush valve. This creates a flow that pushes waste through the trapway, aided by the bowl's design, to achieve siphonic action. The fill valve then replenishes the tank and bowl. The flapper ensures water stays in the tank until the next flush.	When you flush the toilet, a handle lifts a flap at the bottom of the tank where water is stored. This water rushes into the toilet bowl, washing everything down a big pipe. At the same time, the tank fills up with water again, getting ready for the next flush. There's also a special float in the tank that stops the water at just the right level.

Table 14: Audience-specific responses from GPT-4

Second, defining the effect of the *audience* or *community* on the LLM during the use-phase, requires a progression into the second stage of Glăveanu's "*developmental trajectory of creativity*" that describes the two stages of "*first becoming able to observe and make use of affordances in the surrounding environment and then mastering this use and altering affordances, adapting what already exists and creating new artifacts with new affordances*" (2013, p. 76). This second stage inherently demands the ability to learn from the observation of new propositions. As described in the 'mental iteration design model', as established earlier in this thesis, to design is to learn from propositions. LLMs, however, are currently only capable of interpreting and proposing ideas, while not being able to recompose their fundamental knowledge structure (or set of instantiated schemata). Franceschelli & Musolesi (2023) make a similar remark, stating that "*the current generation of LLMs are immutable entities, i.e., once the training is finished, they remain frozen reflecting a specific state of the domain*" (p. 6). Although capable of interacting with the language based on the extracted features that underpin it, current LLMs can be compared to a human, aged 30, that can only act based on the knowledge that has accumulated during events up to, for example, its 20<sup>th</sup> year. Input prompts, textual accounts that become part of the context of the LLM, allow the LLM to construct new combinations of tokens, but there is no mechanism that allows the LLM to transform its underlying structure of logic through a process of reflecting on the differences between the expected and interpreted artifact.

Instead of learning during the use phase, the LLM can only alter the direction of its output by considering the user-input and its own generated content in its context-window. The context-window is the amount of tokens that an LLM can attend to in its prediction for next most-probable token (OpenAI et al., 2024). An analogy between the context-window of LLMs is the 'Random Access Memory' (RAM) in current computer systems<sup>12</sup>. As such, the *context* of an LLM after training can be compared to its context-window.

#### Takeaway box 11

The set of social and material components that constitute the *press* of an LLM during training are its textual training dataset (its inspiring set (Ritchie, 2007)) and human feedback (during the process of RLHF). After training, the LLM's *affordances* or *context* are the tokens in its context-window. As such, LLMs cannot learn during the use phase. The LLM's context-window, however, does enable it to adapt to the user's input and to thus co-shape the direction of the design process. This process of continuous adaptation is however limited by the size of the context window of the LLM.

---

<sup>12</sup> This observation is presented by Karpathy (2023), a leading AI researcher, in a video on his youtube channel: [https://www.youtube.com/watch?v=zjkBMFhNj\\_g&t=2536s](https://www.youtube.com/watch?v=zjkBMFhNj_g&t=2536s)

### 7.2.3 Hypotheses on the LLM's design-augmentation capabilities

Relating to the questions stated at the beginning of this section, this section aims to project the findings about LLMs relative to the *person* and *press* perspectives on creativity onto a set of hypotheses about their ability to augment a designer by viewing LLMs in the light of the *process* and *product* perspectives of creativity literature. First both perspectives will be considered in the assessment of the *novelty*-related questions, afterwards to address the *quality*-related questions, and finally the question surrounding *directionality*.

#### *Novelty*

This subsection aims to address the following questions:

1. Are LLMs capable of introducing knowledge components and propositions to increase the novelty of the designer's output?
2. Are LLMs capable of producing novel concept propositions relative to the context in which it is externalized?

Regarding the *process*-related creativity of LLMs, Franceschelli and Musolesi (2023) argue that LLMs are at most capable of what Boden (2009) calls *combinational creativity*. Recalling the visual in figure 26, this implies working with existing elements and finding new combinations between them (Boden, 2009). As such, the answer to the first question of this subsection would be that LLMs are not capable of introducing novel knowledge components, since it can only work with those that are part of its knowledge base or introduced by the user. The answer to the second question then is that LLMs are capable of combining these knowledge components into novel concept propositions. However, although these answers seem logical successions to the statement of Franceschelli and Musolesi (2023), there are some nuances that have to be added to the discussion when not only the LLM, but the LLM relative to a specific designer is concerned.

First, an important distinction has to be made between the absolute novelty of the proposition and the subjective interpretation of the designer that defines whether the absolute level of novelty is identified and/or realized. Regarding the absolute novelty of a proposition, consider the following: LLM-based combinational creativity can be expected to remain within the boundaries of what is represented within the training data, due to the fixed established relationships between tokens after training. In other words, no combinations will be presented that are new to the history of humanity, although they might be new to the receiving individual. Boden (2009) calls the former H-novelty and the latter 'P-novelty', where the 'H' stands for historical and 'P' for 'personal'. As such, some might interpret the results of the LLM as being novel, while others might find them to be trivial. Therefore, some level of subjectivity is present when P-novelty is the metric of investigation, while H-novelty can be more easily checked by comparing against the content in the training dataset and other sources of information.

Let us now consider this subjective interpretation in more detail. Given the scenario of a knowledge component or concept proposition that is not H-novel but that has the potential to become P-novel (it still needs to be embedded in a dominating schema to form a conceptual *structure* that might lead to the desired *behavior*), the implication is that the LLM should provide a logical line of reasoning that connects the design brief or presented problem or *function* to the presented knowledge component, such that the individual is actually capable of interpreting the content in the correct manner. A lack of visible reasoning might lead the user to discard the presented knowledge domains or consider the propositions not to be novel. For example, recalling the factor of *paradigm relatedness* that underpins *novelty* (see figure 26 and slightly adapted in figure 37), if the LLM were to introduce a knowledge component from outside the personal knowledge paradigm (visualized as the smaller dotted square) without providing an explanation of its relation to the other components within it, it depends on the way in which the receiving designer handles that proposition whether that knowledge is embedded into his set of schemata to form a novel proposition. If, however, the LLM also provides an explanation of the link between the knowledge presented (with novelty potential once correctly interpreted) and the design task, the potential novelty may be transformed into actual novelty as the designer is able to connect the idea to its own frame of reference.

As mentioned before, a hybrid form of bottom-up and top-down augmentation can be used, with the former being materialized through an analogy that translates the generated content to the knowledge domains of the designer. Expressed in terms of schemata, the designer is then enabled to embed the instantiated schema (an instance of *structure*) introduced by the LLM to the dominating schema that represents his conceptual understanding. Based on this observation, one can argue that true novelty still finds its origin in the human's brain, although the potential for such novelty to emerge can be increased by an LLM's proposition of knowledge components and their connection to the design problem beyond the imagination of the human's brain.

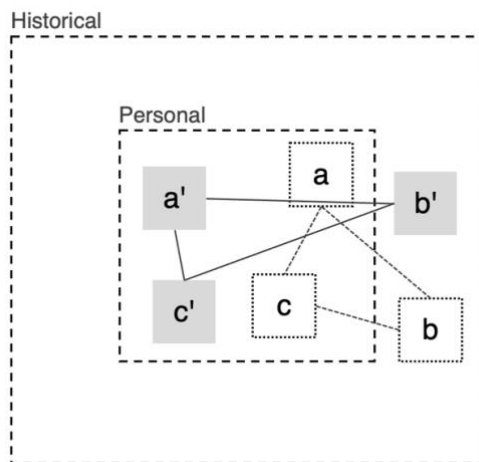


Figure 37: Combinational creativity with P- and H-novelty boundaries

The question then arises as to whether LLMs are capable of producing connected lines of reasoning between their context-relative ‘new’ knowledge domains and propositions and the design task or brief at hand. A type of reasoning that enables humans to solve problems through new propositions based on prior knowledge is analogical reasoning (Qin et al., 2024). Analogical reasoning is a type relational reasoning that allows humans to bridge various domains of knowledge by finding common underlying logic or composition (Vosniadou & Ortony, 1989; Weinberger et al., 2016). Recalling from earlier in this thesis, analogous reasoning is the use of *design patterns* that transfer content about *structure* (as defined in the FBS ontology) from source to target domain (Goel & Bhatta, 2004). Analogies form the chain of causally related knowledge components (*structure*) that explains the link between *function* and *behavior*. Since LLMs are approaching human-level analogical reasoning (Musker et al., 2024), they should in theory be able to support the human in interpreting the proposition using this type of reasoning as a fundament for explanations about the novelty and logic of their propositions considering the context in which they are expressed.

Second, considering the answers to both questions, a distinction has to be made between LLMs that are simply dreaming their training data and LLMs that are prompted by a user. Considering the former, this is when LLMs generate language without a specific input request from a user<sup>13</sup>, as such creating strings of tokens that are most likely to follow up on each other given the training dataset. Here, LLMs cannot be expected to generate H-novel content as, by default, these models are inclined to represent the average distribution of their data (Franceschelli & Musolesi, 2023). However, as discussed in the previous sections, prompts can change this. First of all, LLMs can identify the level of semantic distance given a certain input text (Organisciak et al., 2023). As such, prompting an LLM to generate an idea that is semantically distant from the domain of the design task can result in paradigm unrelated ideas and hence increase novelty. Furthermore, user-prompts, often being combinations of tokens that are not exactly part of the LLM’s training dataset, can alter the process of language generation since the decoder stack includes a combination of Encoder-Decoder and Masked self-attention layers that ensure that the generated language attends to both the input prompt as well as the semantic and syntactic features of the rest of the training dataset. The mechanism of next-word prediction based on existing data therefore might therefore not be as big of an obstacle to novelty as intuition would suggest.

In theory, the implication is that both the novelty of the proposition given the context in which it is expressed as well as the novelty of the combination of knowledge components can be positively affected by user prompts, as they influence the way in which the LLM represents knowledge components through natural language. Although

---

<sup>13</sup> This observation is presented by Karpathy (2023) in this video:  
[https://www.youtube.com/watch?v=zjkBMFhNj\\_g&t=2536s](https://www.youtube.com/watch?v=zjkBMFhNj_g&t=2536s)

proposed tentatively, one could argue that this can lead to H-novel propositions. Although the H-novel proposition does not completely originate from the LLM as it requires an input-prompt crafted by the user, the LLM allows propositions to emerge from the words of the designer by extending them into domains of knowledge that might not have been considered by the designer, or any other human for that matter.

As such, regarding the first question, LLMs (including user-prompts) should be able to generate P-novel knowledge domains, or accounts of *structure*, including explanations about their relation to the definition of the *function* that the design process aims to realize in the form of analogies. Considering the second question, H-novelty is least probable, but not impossible, while the P-novelty of either knowledge or concepts depends on the individual receiving and interpreting it. Both the knowledge and concept forms of novelty augmentation depend on the interpretation of the user and as such benefit from a translation to the known knowledge domains of the individual. As such, the statement from Franceschelli and Musolesi (2023) still holds, although one can argue that also the novel combinations of existing knowledge components can only be materialized once a user-generated prompt is involved and the user is correctly educated on the novelty potential of the LLM-generated content.

The fact that H-novelty is least probable should not be considered a big negative to the ability of LLMs to augment the novelty of the designer's work. After all, creativity literature highlights the cumulative and historical nature of creative artifacts. It could even be argued that it is impossible for creative artifacts to fully break from the past, as they are grounded in the conceptual thinking based on words and notions that represent the 'cultured' experiences from the past (Weisberg, 1993). In the words of Barron (1995):

*"The human act of creation, basically, is a personal reshaping of given materials, whether physical or mental. What is new is form transformed; a new form, generated from an old one"* (p. 313)

Vygotsky proposes a similar argument, arguing that *"existing tools and symbols are fossilized thoughts and ideas of people who have come before in history"* (Moran & John-Steiner, 2003, p. 79). Following this same line of reasoning, Kantosalo and Takala refer to *artifacts* as being *contributions* in the larger collaborative act (2020). This line of reasoning also points at the cumulative character of creativity – creativity is collaborating with *artifacts* that resulted from countless historical human interactions, where *artifacts* or *products* are *contributions* (Kantosalo & Takala, 2020) in a larger web of ideas that together form one's conceptual understanding. Considering the objective here, designer augmentation, small steps of P-novelty could be the steppingstones towards H-novel solutions that would potentially not have been existed without the many P-novel domains of knowledge and concepts realized through the LLM-designer interaction. As such, the following hypothesis can be constructed:

**Hypothesis 1:**

When provided with the right user-prompt, and when being informed on the knowledge domains known to the designer, LLMs are capable of generating language that contributes to the novelty of the ideas or concepts as perceived and conceived by the designer.

**Takeaway box 12**

H-novel creativity still originates from the human brain, whereas LLMs exhibit combinational creativity rather than generating truly novel ideas (H-novel). However, in the context of designer augmentation novelty is relative. Here, an LLM can enhance the novelty of the designer's output when it provides sufficient explanation on its reasoning behind a certain proposition and when being prompted in the right way.

*Quality*

This subsection aims to address the following question:

3. Are LLMs able to evaluate the workability, relevancy, and specificity of the total of design ideas, referred to as the quality of the solution?
4. Are LLMs capable of producing a quality (workable, relevant, and specific) and logically structure proposition based on its own knowledge and content shared by the designer?

Although an LLM might be able to generate P-novel content (following the definition of P-creativity from Boden (2009)), this does not mean that it is useful given the task at hand. The mechanism that ensures that the LLM's generated content addresses the questions or concerns presented in the user-prompt, the LLM's training phase also includes a finetuning phase using human feedback, called 'Reinforcement Learning from Human Feedback (RLHF)'. One could argue that this is the mechanism that is related to the *quality* of the output, as the LLM is trained to respond in a manner that is similar to how humans prefer to converse and exchange information. Although this partially addresses the definition of *quality*, as combination of *workability*, *relevance* and *specificity* (Sääksjärvi & Gonçalves, 2018), it is necessary to understand the ability of LLMs to spot gaps in the chain of one's logic or reasoning connecting the status quo and desired behavior. How can an LLM's mechanism of language creation account for this?

One way to go about this, is to prompt the LLM to be critical and spot gaps in one's proposition relative to the overall design goal. The word-embeddings of 'critical' and 'gaps' combined with the mechanism of next-word prediction may trigger the chains of logic captured in these vector-representations and transfer that logic to the specific scenario of the user. For example, the word 'critical' may be stored in the LLM as a word

that is positioned closely to words such as ‘feedback’ and ‘critique’. These words can be positioned in hypothetical training data sentences such as:

- “I am slightly critical towards his rhetoric, as he does not seem to be able to manifest what he claims he is capable of”
- “A point of feedback would be to slightly re-orient your current direction and look for ways to include expertise from other domains as well.
- “My only critique is that the author did not integrate storyline A and B well

During training the LLM constantly takes into account the words surrounding the word of interest in order to establish its vector-embedding (Kalyan, 2024). As such, the syntactical structures of sentences in which the word ‘critical’ occurs, co-inform the semantic and syntactical structures that are generated due to that word occurring in a user-prompt.

Another approach is to prompt an LLM to execute a semantic similarity test by comparing the input from the user to the solution it would generate itself based on the knowledge stored inside the model (Banjade et al., 2024). Using the method from Banjade et al. where the LLM is prompted to spot gaps in student explanations of their code (2024) but translating it to design process augmentation, the LLM is first prompted to generate a set of reference knowledge components or ideas. Secondly, it is prompted to perform a semantic similarity test between its own generated content and that produced by the designer using the prompt shown in figure 38. Research identified that using this approach over a prompting strategy that directly commands the LLM to find gaps, increases the correctly identified gaps by around 34%, measured across four state-of-the-art LLMs (Banjade et al., 2024).

Given the following design task:{design task} and the following reference knowledge domains/ideas: { knowledge domains/ideas }, your task is to identify what is missing in the following designer’s knowledge domains/ideas:{designer’s knowledge domains/ideas } of the design task. Do so by providing a semantic similarity score on a scale of 0 to 1, 0 being least similar and 1 being most similar, for the following knowledge domains/ideas: {reference knowledge domains/ideas} and {designer’s knowledge domains/ideas }.

Figure 38: adapted from Banjade et al. (2024)

An observation from the novelty augmentation capability section that is relevant here as well, is that it depends on the individual designer and its expertise in certain knowledge domains whether the LLM is capable of augmenting him. Nonetheless, given the observations in the previous and this subsection, the following hypothesis can be conceived:

### **Hypothesis 2:**

When provided with the right user-prompt, and when being informed on the knowledge domains known to the designer, LLMs are capable of evaluating the



*workability, relevance and specificity* of design ideas and concepts, and to propose the missing critical domains.

### Takeaway box 13

Similar to novelty augmentation, the quality enhancements of the LLM's propositions are relative to the propositions of the designer and require prompting. The quality augmentation can be realized by prompting the LLM to be critical and spot gaps in one's proposition relative to the overall design goal. Another approach is to prompt an LLM to execute a semantic similarity test by comparing the input from the user to the solution it would generate itself based on the knowledge stored inside the model.

Because the method proposed by Banjade et al. (2024) is used, this also implies that the LLM can generate a logically structured proposition itself. As such hypothesis two covers both question three and four.

### Directional

Concluding from the sections about *person*, *press*, and *product*, an LLM does not have the characteristics to initiate a creative process from itself, it is not capable of learning during its use-phase, but there is the potential for both user-prompts as well as LLM-generated content to simulate these characteristics and to 'adapt' the direction of the output by considering previously generated tokens. The final component of this enumeration highlights the dual nature of the *product* and is the bridge between accounts of LLM novelty and quality from the previous two subsections and this subsection that is interested in the ability of LLMs to generate content in an iterative manner. This subsection aims to address the following question:

5. Can LLMs coherently progress through an iterative process in which it enables the designer to arrive at new insights?

First, let's consider the dual nature of the *product*, or the output. The *product* is both the outcome of a process – more literally, it is the 'product' of a *process*, as well as a part of the chain of actions that constitute the creative process itself. This highlights the intertwined nature of *product* and *process* again; the *product* is the aggregate of *process*-related insights and characteristics, whereas the *process* can be expressed in terms of the iterations of the *product*. Reflecting on these contemplations, one can argue that what is considered the *product* is as much material to the *process* as it is the 'end state' of the creative act. As such, building on Glăveanu's notion of the cumulative nature of creativity (2013) (Barron, 1995; Sawyer et al., 2003) and the fact that an *artifact* can be classified as an *action* (Sawyer, 1997)(Cropley, 2006), the language generated by an LLM is both the LLM's action taken to invite its *audience* to give meaning to it, as well as a steppingstone within its own creative *process* of meaning-making. This description fits

well with the Geneplore descriptive framework that describes the creative process as a combination of generative processes that produce candidate ideas, and exploratory processes that elaborate on the inherent potential of these candidate ideas (Kaufman & Sternberg, 2010, p. 94). Besides, it also corresponds to the design process framework as introduced in section five of this thesis, where the *proposition* stage can be regarded as a generative process and the *interpretation* and *evaluation* as explorative processes.

This idea of *proposition* and *interpretation* opens new ways to approach the creativity of the LLM as not only their mechanism of creating natural language, but also the language they create is process-related material. This is an extension of the insight presented in the previous subsection, which states that the user prompts form an important part of the process that leads to creative outcomes. Furthermore, it strengthens the statement that the mechanism of next-word prediction might therefore not be as big of an obstacle to novelty and thus creativity as intuition would suggest. Referring to Jay and Perkins (1998), this insight changes the way in which LLMs should be understood in terms of their creative capacities. The LLM's output should be regarded as a *contribution* (Kantosalo & Takala, 2020) to the *press* in which it operates, with the *press*, or *context*, being the context window in which the LLM generates language. The context-window of the model can be regarded as its temporary playground for manipulating its output while interacting with the initial prompt from the user.

A conceptual parallel could be made to 'system one' and 'system two' thinking, a concept that emerged in the field of research about human reasoning (Evans, 2003). System one consists of a group of independent subsystems that encompass domain-specific knowledge obtained through a domain-general learning process (Evans, 2003). System two enables abstract reasoning and hypothetical thinking. It is limited by the capacity of working memory and is associated with general intelligence metrics (Evans, 2003). Translating this to the domain of LLMs, the generation of language could be regarded as a 'system one' activity as the model can autonomously perform this action. The ability to reflect and make iterations then is part of 'system two', where a human gets involved to enable the LLM to follow a certain process-related logic through a carefully constructed prompt. The prompting technique closely related to this is 'chain of thought' (CoT) prompting where the user prompts the LLM to take intermediary steps to solve the question or problem<sup>14</sup>.

There are, however, three limiting factors when considering LLMs in the context of this iterative process. First, as mentioned before, the LLM is not capable of updating its fundamental structure because of the reflection on its output. It is only able to expand the domain of tokens associated to the domain of interest; the LLM has

---

<sup>14</sup> As further explained on this webpage: <https://www.metadocs.co/2024/09/17/how-openai-o1-works-in-a-simple-way-and-why-it-matters-for-rag-and-agentic/>

generated new tokens that were not part of the user-prompt and hence triggers tokens that were previously not connected. The second limiting factor is the context window of the LLM, defined as the set of tokens that co-predict the next set of tokens (Kalyan, 2024), as after a certain amount of tokens, older tokens are not taken into account anymore. Re-iterating the analogy from Karpathy (2023), the LLM's context window can be regarded as its Random Access Memory (RAM). Finally, although recent research aimed to boost the 'reasoning' capability of LLMs with techniques including Chain-of-Thought (CoT), empirical research on the 'reasoning' capabilities of LLMs highlights a discernable gap in the reasoning performance between various LLMs (GPT-4, GPT-3.5-turbo and GPT-3.5-Instruct) and humans, especially when confronted with longer, more complex or disjunctive forms of logic (Wang et al., 2024).

Iteration in the context of LLMs is thus dependent on the prompt that instructs the LLM to perform such an iterative process and limited by the LLM's inability to learn after training, context-window and reasoning abilities. If the iteration were to demand the process of learning, then LLMs would not be able to perform an iterative process. However, iteration, defined in the Cambridge dictionary as *"the process of doing something again and again, usually to improve it, or one of the times you do it"*, does not require learning as it describes a process of revision on a descriptive level. As such, LLMs, although not capable of altering their fundamental structure, can participate in an iterative process, where generated tokens trigger new tokens.

Leveraging the insight that the LLM-generated content is also process-related material, two observations can be made. First, a distinction between continuous and interrupted iteration has to be established. Continuous iteration occurs when an LLM is prompted to iteratively generate language. That is, it is instructed to follow a sequence of steps that form a cyclical process in which after each cycle the LLM reflects on the output and either confirms or alters the direction of the output itself. This can be done by either altering the premises, the ideas, or their positioning relative to the knowledge structure – the composition of the concept, or the overall *structure* that leads to the desired *behavior*. Interrupted iteration is when the designer and LLM alternately perform a 'mental iteration' (*understanding* and *proposition*). Here, the human can complement the LLM's inability to learn and limited reasoning abilities and vice versa. As such, for augmentation purposes this is the preferred *modus operandi*. Relating back to question five, and combining the insights from this section with those of the sections about the novelty and quality of the LLM's output, the following hypothesis can be established:

### **Hypothesis 3:**

LLMs can coherently progress through an iterative process, while both addressing the designer's input and emerging insights, and enabling the designer to get to new insights, within the boundaries of its context window.

**Takeaway box 14**

The language generated by an LLM is both the LLM's action taken to invite its *audience* to give meaning to it, as well as a steppingstone within its own creative *process* of meaning-making. This idea of *proposition* and *interpretation* opens new ways to approach the creativity of the LLM as not only their mechanism of creating natural language, but also the language they create is process-related material. A conceptual parallel could be made to 'system one' and 'system two' thinking. The limitations of this approach include 1) the fact that the LLM cannot alter its fundamental knowledge structure during generation, 2) the context window size of an LLM, and 3) the reasoning capabilities of the current generation of LLMs.

!

It should be noted that the newly introduced OpenAI o1 LLM<sup>15</sup>, integrates a lot of observations presented in this section. This model has finetuned itself based on the most successful variants of self-generated chains-of-thought (CoT). Several sequences of chains-of-thought (CoT) are generated and evaluated on their contribution to the final prediction of the sequence of generated text (Zelikman et al., 2024). The model is then finetuned based on those CoT sequences that contributed most and are thus the most powerful. As such, it has learned what lines of reasoning present the best outcomes given a certain input prompt. This new training paradigm mostly affects the identified limitations in the domain of reasoning. The other two limitations still remain.

Second, continuous iteration holds promising potential for increasing the novelty and quality of the LLMs output, and hence the success of the augmentation of the designer. With regards to *quality*, one can imagine the LLM being capable of 'reasoning' (to a limited extent (Wang et al., 2024)) its way towards the gap in either *completeness*, *relevancy*, or *workability* by constructing logically related propositions. This process has to be initiated by a prompt that allows the LLM to generate such a 'chain of thought' (CoT). Considering *novelty*, the ability to expand the scope tokens that form the fundament for further generation has the potential to trigger a set of tokens from a more diverse range of knowledge domains. This might increase the novelty of the *concept* and hence the combinational creativity of the LLM's propositions (Boden, 2009). Building on the theory and hypotheses presented in this and the previous three subsections, the fourth hypothesis can be conceived:

**Hypothesis 4:**

Continuous iterative generation can improve the novelty and quality of the LLM's output.

---

<sup>15</sup> <https://openai.com/index/introducing-openai-o1-preview/>

!

This final hypothesis can be regarded as partially validated by the release of the newest LLM from OpenAI, o1, which takes this mechanism and uses it to finetune the LLM<sup>16</sup>. During training, this training mechanism takes away the limitation that LLMs, when in use, cannot alter their knowledge structure. During use, it uses the most successful line of reasoning to generate a framework for answering the user's question.

#### 7.4 A definition of LLM-based (& human) behavior

Based on the technical view on the mechanisms of LLMs (section 7.2) and the reflection on this mechanism in the light of creativity behavior (section 7.3), the following description of the behavior of an LLM-based collaborative actor can be provided: Initiated by the tokens in the LLM's context window that comprise descriptions of both design content and content from which can be derived the inferred state of the Human, supported by its knowledge, mediated by its internal state (tokens in context window), steered by its goals (tokens in context window), and constrained by its attributes (both from its own mechanism as well as the tokens in context window), the LLM takes action. Table 15 embeds the LLM-related characteristics related to each actor-related set of factors as described in the framework from Viros-i-Martin and Selva (2021) to complete the translation of human-AI design space exploration to LLM-augmented design.

Human-AI collaborative design space exploration framework				Mental iteration design process & LLM capabilities
Perception	Estimated state of the world	Human	An approximation of the state of the design task	Part of the <b>understanding</b> , relating to the <b>concept/memory space</b> stage of a mental iteration.
		CA (LLM-based AS)	[Not defined, implied to be similar to the human]	An approximation of the desired command of a user based on the tokens within the context window of the LLM.
	Inferred state of the CA/human	Human	The expected interpreted state of the CA given the interaction	The choice for a persona determines the state of the CA

<sup>16</sup> As further explained on this webpage: <https://www.metadocs.co/2024/09/17/how-openai-o1-works-in-a-simple-way-and-why-it-matters-for-rag-and-agentic/>

		CA (LLM-based AS)	[Not defined, implied to be similar to the human]	The choice for a persona determines the state of the human
Factors that constitute the state of the agents	Internal state (Dynamic during design process)	Human	Motivation, cognitive workload, situational awareness, <i>trust in CA</i> , <u>design preferences</u> , the <u>degree of design fixation</u>	Defined by the level of perceived <b>epistemic uncertainty</b> and as such dependent on the design task <b>relative to the designer's knowledge</b>
		CA (LLM-based AS)	Dialogue state, proactivity, role ('historian', analyst, explorer, expert, critic)	The dialogue state is based on the <b>tokens within the context window</b> that define the dialogue state, the proactivity and the role of the LLM. The proactivity and role depend on how the LLM is prompted. <b>This internal state can be compared to the type of augmentation the LLM is supposed to materialize.</b>
	Knowledge	Human	Domain, design process, problem specific	Relating to <b>knowledge space</b> in the design process framework, which is <b>dynamic</b> and extendable
		CA (LLM-based AS)	Problem database, heuristic database, historical designs	Relating to the embedded form of the LLM's training dataset, which is static.
	Goals	Human	Design, learning, collaboration	A designer's goals are positioned within the 'activity' layer of a designer's behavior. The <b>learning</b> goal is to alleviate one's perceived uncertainty through the proposition of candidate ideas. This is aligned with the design goal. The <b>design goal</b> follows the same mechanism <b>[collaboration goals]</b>

		CA (LLM-based AS)	Design, learning, collaboration (alleviating design fixation, increasing common ground and trust in CA)	The LLMs <b>goal is to augment the designer</b> by delivering the novel knowledge or idea propositions, while also ensuring the quality of the concept development. As such, <b>it is aligned with the designer's goals</b> . In doing so, it alleviates design fixation. [Increasing common ground and the designer's trust in LLM]
	Attributes (Static during design task)	Human	Cognitive style, level of expertise, risk aversion, expectations, preconceptions	[Person-based attributes]
		CA (LLM-based AS)	Learning, intelligence, autonomy, communication	LLMs <b>cannot learn</b> (learning), have <b>some intelligence</b> as they are able to process information to generate new information (intelligence), have a <b>varying level of autonomy</b> depending on how they are prompted and embedded (autonomy), and have <b>excellent natural language processing capabilities</b> (communication)
<p>Legend</p> <ul style="list-style-type: none"><li>- <i>Italic</i> = Related to human machine collaboration literature*</li><li>- <u>Underlined</u> = Related to design literature*</li><li>- Orange = Not covered so far by prior literature in this thesis.</li></ul> <p>*(Viros-i-Martin &amp; Selva, 2021)</p>				

Table 15: The design process framework and LLM-related capabilities related to the framework for human-AI collaborative design space exploration (Viros-i-Martin & Selva, 2021)

## 7.5 Practical approach: LLM augmentation novelty, quality & iteration

This section reflects on the validity of the first two hypotheses conceived in the previous section through practical examples. These first two are the most important given the focus on their novelty and quality enhancement potential. The third and fourth hypothesis are more complex and as such require more extensive analyses in future research. These two practical examples provided in this section give a basic understanding of the validity of the first two hypotheses but cannot exhaustively confirm their validity. Nonetheless, they are a useful vehicle for further interpretation and a pointer for future research into the design process augmentation capabilities of LLMs.

### 7.5.1 Hypothesis 1

Staying close to our example of a window-design, the prompt shown in table 16 was given to GPT-4o. Table 16 also shows the response from GPT-4o.

Input prompt novelty
I want to design inlets for light positioned in the wall of a building that make sure that the heat in summer stays out while still enabling cool air to flow in. Come up with an idea that is semantically distant from what I've mentioned before. Then come up with an analogy that connects that idea that is semantically disconnected from the design project of designing the inlet for light and fresh air that I mentioned at the beginning.
GPT-4o's response
<p><b>Semantically Distant Idea</b></p> <p>Consider designing a biophilic element that uses plants to regulate temperature and airflow. A vertical garden system can be incorporated into the wall of a building, where plants act as a natural insulator. The foliage can provide shade, reducing heat gain during summer, while the transpiration process cools the air.</p> <p><b>Analogy</b></p> <p>Think of this system as a forest canopy. In a forest, the tree canopy shields the ground from direct sunlight, maintaining cooler temperatures and allowing for a comfortable airflow. Just as the tree leaves create a microclimate that supports diverse life forms, the vertical garden on the building wall would create a microenvironment that regulates temperature and airflow, enhancing the comfort and aesthetics of the space.</p> <p><b>Connection to Initial Idea</b></p> <p>In this analogy, the forest canopy is equivalent to the vertical garden system. Both the tree leaves and the plants in the vertical garden serve as natural filters and insulators. They provide shade to keep the interior cool while allowing air to flow through, similar to the way you want the inlets for light and fresh air to function. The semantically distant idea of using plants integrates nature into the built environment to achieve the desired thermal and airflow regulation, paralleling the functional goal of your original design project.</p>

Table 16: Hypothesis one test



Two things stand out. First, reflecting myself on the answer generated by GPT-4o, it can be said that the idea is neither P-novel or H-novel, as I have come across such a solution before. However, it might be novel solution considering the more specific circumstances in which this project might be executed. For example, if this building were to be built for social housing, and it would use the presented type of natural cooling system while being built at a very low cost, it could be considered a novel solution. Again, this shows the high dependency of novelty on contextual factors. Second, GPT-4o was able to generate a fitting analogy that explains in simple terms how the solution connects to the design brief, such that I, as the user or designer, am able to interpret the solution with the knowledge at my disposal. In general, one could argue that the LLM wouldn't have generated this form of answer when not being prompted in this specific way. However, if augmentation is the objective instead of defining whether LLMs are creative entities, this does not stand in the way of accepting the first hypothesis given the limited validity of this exploratory example.

7.5.2 Hypothesis 2

Building on the input prompt and response of the previous subsection, this subsection shows the LLM's performance in terms of quality augmentation. Two different approaches are tested. First, the method of prompting the LLM to be critical and spot gaps directly is used (table 17). Second, the method from Banjade et al. (2024) for the analysis of gaps in the formulation of an explanation, or in this case the formulation of a design idea, is used for the prompt design (table 18).

Input prompt quality 1
I want to design inlets for light positioned in the wall of a building that make sure that the heat in summer stays out while still enabling cool air to flow in.
I have currently conceived the following idea: A vertical garden system can be incorporated into the wall of a building, where plants act as a natural insulator. The foliage can provide shade, reducing heat gain during summer, while the transpiration process cools the air.
Be critical and identify gaps in the ability of my idea to realize my design objective.
GPT-4o's response
Your idea of incorporating a vertical garden system into the wall of a building to act as a natural insulator and provide shade, reducing heat gain while cooling the air through transpiration, is innovative and sustainable. However, there are several considerations and potential gaps to address to ensure it effectively meets your design objectives:
1. Thermal Insulation Efficiency
<b>Variation in Plant Coverage:</b> The density and type of foliage can significantly affect the shading and insulation properties. Sparse or inconsistent plant coverage might not provide sufficient shade.

**Seasonal Changes:** The effectiveness of the vertical garden might vary with seasons. Plants might lose leaves in colder months, reducing insulation and shading.

**Maintenance:** Regular maintenance is required to keep plants healthy and ensure consistent coverage. Dead or unhealthy plants can reduce the insulation effect.

## 2. Airflow Management

**Passive Cooling vs. Airflow:** While transpiration can cool the air, it might not necessarily direct cool air into the building. There needs to be a mechanism to ensure that the cooled air enters the building effectively.

**Air Circulation:** Simply having cool air outside does not guarantee it will flow into the building. There needs to be a design that promotes natural ventilation, such as strategically placed vents or openings that facilitate airflow.

## 3. Humidity Control

**Increased Humidity:** Transpiration increases humidity levels. High humidity can lead to discomfort and potential mold growth inside the building if not properly managed.

**Condensation Issues:** The increased humidity near the walls could lead to condensation, potentially damaging the building structure over time.

## 4. Structural Integration

**Weight and Support:** A vertical garden system can add significant weight to the wall. The building structure must be capable of supporting this additional load.

**Waterproofing and Drainage:** Proper waterproofing and drainage systems are essential to prevent water damage to the building facade.

## 5. Climate and Location Suitability

**Local Climate:** The effectiveness of vertical gardens can vary with local climate conditions. In very hot and dry climates, the water requirements for the plants might be unsustainable.

**Orientation:** The orientation of the wall (north, south, east, west) will affect the amount of sunlight the plants receive and, consequently, their ability to provide shade and cool the air.

## 6. Energy Consumption

**Watering Requirements:** Ensuring the plants receive adequate water, especially in dry seasons, might require significant energy and water resources, potentially offsetting some of the environmental benefits.

### Recommendations to Address Gaps:

**Optimize Plant Selection:** Use a mix of plant species that provide dense foliage year-round and are well-suited to the local climate.

**Design for Ventilation:** Incorporate passive ventilation design principles, such as cross-ventilation, to ensure cool air flows into the building.

**Humidity Control:** Implement measures to manage indoor humidity, such as dehumidifiers or enhanced ventilation systems.

**Structural Considerations:** Ensure the building can support the additional weight and has proper waterproofing and drainage systems.

**Regular Maintenance Plan:** Develop a maintenance plan to keep the vertical garden healthy and functional.

**Smart Watering Systems:** Use efficient irrigation systems, such as drip irrigation, to minimize water usage.

By addressing these gaps, you can enhance the effectiveness of your vertical garden system in achieving the dual objectives of reducing heat gain and promoting cool air inflow.

Table 17: Hypothesis two test version one

#### Input prompt quality 2

Given the following design task:{I want to design inlets for light positioned in the wall of a building that make sure that the heat in summer stays out while still enabling cool air to flow in} and the following reference knowledge domains/ideas: {Consider designing a biophilic element that uses plants to regulate temperature and airflow. A vertical garden system can be incorporated into the wall of a building, where plants act as a natural insulator. The foliage can provide shade, reducing heat gain during summer, while the transpiration process cools the air}, your task is to identify what is missing in the following designer's knowledge domains/ideas:{A vertical garden system can be incorporated into the wall of a building, where plants act as a natural insulator. The foliage can provide shade, reducing heat gain during summer, while the transpiration process cools the air} of the design task. Do so by providing a semantic similarity score on a scale of 0 to 1, 0 being least similar and 1 being most similar, for the following knowledge domains/ideas: {reference knowledge domains/ideas} and {designer's knowledge domains/ideas }. Then generate the missing components in the designer's idea.

#### GPT-4o's response

##### Semantic Similarity Score

To calculate the semantic similarity between the reference knowledge domains/ideas and the designer's knowledge domains/ideas, let's first analyze the key concepts and ideas in both sets.

##### Reference Knowledge Domains/Ideas:

- Biophilic design
- Plants regulating temperature and airflow
- Vertical garden system
- Natural insulation
- Foliage providing shade
- Transpiration cooling the air

##### Designer's Knowledge Domains/Ideas:

- Vertical garden system
- Plants as natural insulators
- Foliage providing shade

- Transpiration cooling the air

### Analysis

Both sets of ideas share common elements such as the vertical garden system, the role of plants in insulation and shade, and the cooling effect of transpiration. However, the reference knowledge domains/ideas emphasize **biophilic design** and the role of **plants in regulating airflow** more explicitly. The designer's ideas do not mention **biophilic design** or how **airflow is specifically managed** through the system.

### Semantic Similarity Score

Given the overlap in key concepts but missing specific elements like biophilic design and airflow regulation, the semantic similarity score would be around **0.85**. This high score reflects substantial similarity, though it acknowledges the absence of some important concepts.

### Missing Components in the Designer's Idea

To align the designer's ideas more closely with the reference knowledge domains, the following components should be incorporated:

**1.Biophilic Design Integration:** The concept of biophilic design, which emphasizes the connection between nature and the built environment, should be explicitly mentioned. The design should focus on not just the practical aspects of temperature control but also on enhancing the occupants' well-being through a connection with nature.

**2.Airflow Management:** The designer should consider and describe how the vertical garden system can be designed to optimize airflow. For example, the arrangement of plants and the design of the inlets can be tailored to ensure that cool air flows into the building efficiently, possibly using the chimney effect or cross-ventilation principles.

**3.Specific Plant Selection:** While plants are mentioned generally, the designer could benefit from specifying types of plants that are particularly effective at providing shade, cooling the air through transpiration, and thriving in vertical garden systems. Additionally, plant selection can consider species that contribute to improved indoor air quality.

By incorporating these elements, the designer's ideas would be more comprehensive and aligned with the reference knowledge domains.

Table 18: Hypothesis two test version two

Reflecting on the outputs generated by the two different prompts, there is one specific thing that becomes apparent. The first approach produces a more diverse range of topics that may be related with the proposed idea, whereas the second approach is more confined by the topics covered in the reference idea. The implication is that the former is more suitable for exploratory quality augmentation in the early phases of

concept development, whereas the latter is more useful when the requirements for the final design are already set in stone and the design idea is evaluated against it. Given the range of topics proposed to improve the original proposition of the designer given both approaches, it can be said that LLMs are suitable for augmenting the quality of the design outcome.

**Takeaway box 15**

Hypothesis one can be considered valid, although its validity is highly context dependent. Hypothesis two can also be considered valid, with different prompting strategies leading to a different focus in the quality augmentation. Both conclusions have to be viewed in the light of the narrow scope of the provided practical examples.

**Hypothesis 1:** When provided with the right user-prompt, and when being informed on the knowledge domains known to the designer, LLMs are capable of evaluating the workability, relevance and specificity of design ideas and concepts, and to propose the missing critical domains.

**Hypothesis 2:** When provided with the right user-prompt, and when being informed on the knowledge domains known to the designer, LLMs are capable of generating language that contributes to the novelty of the ideas or concepts as perceived and conceived by the designer.

## 8. An interface for an LLM-based design process augmentation system

---

So far, design theories have been unified to work towards a conceptual foundation for design process augmentation in the context of a collaborative process. However, for LLMs to practically augment the design process, an interface that realizes the desired interaction – hence, an interaction that can realize the descriptions of design process augmentation in the sequential format described in section 6.8 – with the designer needs to be established. Currently, there is a limited view on the practical application of design process augmentation theory (Berni et al., 2024). Although several examples exist, each one has its own limitations. For example, existing interfaces, such as the LLM-based augmentation tool “Supermind Ideator” (Rick et al., 2023) lack a connection to sophisticated accounts of design theory. Similarly, Ailixr<sup>17</sup>, a tool designed to help designers design by enabling the creating of flow schema’s consisting of several different generative AI models, lacks a theoretical fundament that explains its effect on the design process. This section therefore investigates the translation of the conceptual framework of GenAI-human co-design to a practical implementation, as such attending to the fifth research question:

*How can design process augmentation practices be manifested in practice?*

---

### 8.1 Approach

Given the aim to build both an internally and externally valid conceptual framework, it seems fitting to at least address the lack of an interface for an LLM-augmented design process by translating the conceptual framework into a low-fidelity prototype interface. This interface should not only translate theory to practice, but also act as a boundary object for the technological capabilities of LLMs in design, with the objective to stimulate the discussion and reflection on what can be possible and what should be possible. First an augmentation system architecture will be proposed that acts as the fundamental structure for the translation of this thesis’ conceptual findings about design process

---

<sup>17</sup> The following link leads to the official Ailixr website: <https://ailixr.io.tudelft.nl>

augmentation to an LLM-based Augmentation System (AS). Afterwards, a visual user-interface (UI) is proposed that shows how this architecture can be translated into a practical interface.

## **8.2 Augmentation system (AS) architecture**

LLMs generate content based on input prompts. These prompts influence the way in which the LLM responds. Since this thesis aims to construct an LLM-based AS, it is relevant to understand how the LLM-, and thus, prompt-based AS is able to realize the augmentation forms given the definition of the design process as in figure 28 from section 6.8. First, the layers of the architecture will be defined. Second, the connection between the architectural layers and prompts will be established, leading to a hierarchically structured prompt generation schema.

### **8.2.1 Architectural layers**

This section provides a perspective on the architectural structure behind the operation of the AS by building on the layers of the Autonomous Cognitive Entities (ACE) framework (Shapiro et al., 2023) (see figure 36 section 7.2.1). Interestingly, the six layers of the ACE framework can be related to the three layers of design behavior (as described by Cash & Kreye, 2017), which is relevant as these three layers form the basis for describing the design process (see figure 21, section 5.13). Combining the ACE framework and the three layers of design behavior, the AS can be described by three layers: the top (strategy), middle (tactics), and bottom (execution) layers.

By combining both the ACE framework and layers of design behavior, a clear overview is created of the various layers that establish the final outcome or behavior of the AS. This allows for the definition of the scope of both human and computational control in line with the perspective on human agency in the process of human-AI collaboration as defined in the introduction (table 19). Although the exact scope of control can differ per scenario, it can be considered desirable that the human at least controls the direction of the interaction within both the strategy and tactics layer, whereas in the bottom layer (execution), the LLM-based AS has more freedom in its generation of the output. The human observes and interprets the LLM's actions to then exert control again within the top and middle layers of the interface leading to a process where the freedom of the LLM-based AS is still governed by human oversight.

Given the observation that the layers of a designer's behavior can be equated to the layers of the ACE framework, each architectural layer of the AS can be linked to the content within the layers of the design process framework (figure 28). This is relevant as the descriptions about the design process contain relevant contextual information that should be shared with the LLM-based AS for effective augmentation. Table 20 presents an overview of all design process related content that should be exchanged between designer and augmentation system (AS) per architectural layer. It combines content from the design process framework with other observation made throughout this thesis.

LLM-based augmentation system (AS)		Designer	
ACE framework	AS architectural layers	Human role	Design process framework layers
<b>Aspiration layer</b> (Mission, Values, Morals, Purpose)	Top layer: strategy (Defined by human input)	Control	Activity layer (Motivation-led, goal-oriented)
<b>Global strategy layer</b> (global point of view, long-term thinking)			
<b>Agent model layer</b> (Self-awareness, internal monitoring)	Middle layer: tactics (Derivative of human input)	Control	Task layer (Smaller parts attending to goals)
<b>Executive function layer</b> (Planning, forecasting, resource management)			
<b>Cognitive control layer</b> (Task selection and switching)			
<b>Task prosecution layer</b> (Task execution, output to the real world, success and failure detection)	Bottom layer: execution (Based on own data & generative mechanism)	Execution, observation and interpretation of AS output	Action layer (Proposition, interpretation, evaluation)

Table 19: The relation between the ACE framework, interface architecture, and design process

Interface system layers	Required information for an effective LLM-based Augmentation System (AS)
Top layer Strategy	<p>Design process information</p> <ul style="list-style-type: none"> <li>- Design goals</li> <li>- Learning goals</li> <li>- Collaboration goals</li> </ul> <p>Other observations and findings</p> <ul style="list-style-type: none"> <li>- <b>Expectations:</b> “It seems that for proper collaboration to emerge, there should be upfront alignment on the preconceptions about and expectations of the collaborative act” (section 5.16)</li> <li>- <b>Expertise &amp; experience:</b> “Similar to expectations and preconceptions, for augmentation to succeed, it is also important to align on one’s expertise” (section 5.16). “As such, it is important that the augmentation system understands the designer’s relative strengths, his knowledge domains and expertise, and weaknesses to adjust the way in which it interacts” (section 6.6.3).</li> <li>- <b>Personality:</b> “It might be necessary to share one’s level of risk aversion and cognitive style, although the specific details go beyond the scope of this thesis” (section 5.16).</li> </ul>



Middle layer Tactics	<p>Design process information</p> <ul style="list-style-type: none"> <li>- The desired mental iteration stage to be augmented and hence the choice for a specific <i>augmentation persona</i>.</li> <li>- The desired way in which the forms of augmentation are materialized, and hence the choice for a specific <i>interaction view</i>.</li> </ul> <p>Other observations and findings</p> <ul style="list-style-type: none"> <li>- <b>Limitation of top-down augmentation approach:</b> “A way to bypass the inherent limitation to the top-down approach, is to combine it with the bottom-up approach. Here, concepts grounded in unfamiliar (to the designer) knowledge domains are complemented by the associated ‘knowledge elements’, or ideas, and their relationships with both each other and the design question” (section 6.5).</li> </ul>
Bottom layer Execution	<p>Other observations and findings</p> <ul style="list-style-type: none"> <li>- <b>Augmentation capabilities:</b> “Hypothesis 4: Continuous iterative generation can improve the novelty and quality of the LLM’s output” (section 7.2.3).</li> </ul>

Table 20: Thesis findings structured per architectural interface layer

### 8.2.2 Prompt generation schema

Building on the three layers and their respective findings as described above, a hierarchically structured prompt generation schema can be established that directs the interaction between designer and LLM-based AS at different levels. A visual representation is provided in figure 39. First, the designer can, for example, share his goals and expectations regarding the collaborative process within the top (strategy) layer. Second, within the middle (tactics) layer, the designer (via prompting) can choose persona(s) and interaction view(s). This layer thus describes what form of augmentation is desired given a certain situation and how it should be materialized. Finally, the designer inputs a prompt at the execution level.

The information in the strategy layer can be regarded as a prerequisite for effective augmentation, since this content governs the way in which the system characteristics, or personas, in the middle layer are manifested. Similarly, the content in the middle layer, influences the way in which the user input-prompt in the bottom layer is handled – it is an account of content manipulation relative to the strategy layer’s direction. The aggregate of prompts resulting from the process of defining each layer thus defines how the augmentation system (AS) responds. The output of the AS then forms the input for designer, who can direct his actions towards any of the three layers. The next section provides a perspective on the way in which this hierarchical prompt structure can be intuitively embedded in a user interface.

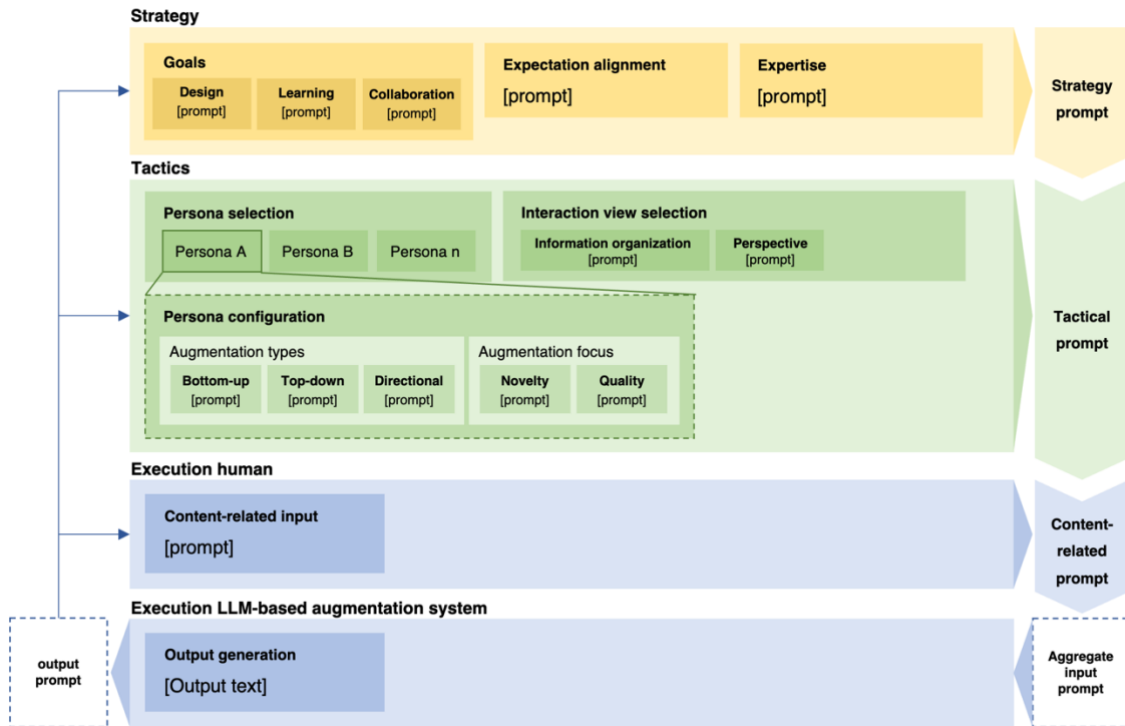


Figure 39: Prompt generation schema

### 8.3 User Interface (UI)

This section translates the conceptual framework of human-AS augmented co-design (figure 28) and the prompt generation schema to an interface that aims to shed light on how the theory about augmentation can be put into practice. The interface uses a digital canvas as working space (figure 40) and builds on the interface fundamentals of tools such as Miro<sup>18</sup> or Figma<sup>19</sup> that replicate whiteboard in the digital domain. This concept is suitable for design as it allows a designer to progress similar to how Crilly (2021b) captures the design process in an idea heterarchy, where ideas about solutions lead to new ideas about problems and vice versa (see figure 5). The digital canvas enables the process of extending one's view by being able to rescale the canvas as the process progresses. Previous design iterations of the interface can be found in Appendix C. The digital canvas can be regarded as the *representation* action through which both *information* and *knowledge-sharing* action (embedded in the *information organization* and *perspective interaction views* respectively) are materialized. Below, an both an interface walkthrough is provided.

<sup>18</sup> <https://miro.com/nl/>

<sup>19</sup> <https://www.figma.com>

!

The interface is a dummy and not functional. If text/visual components are highlighted in purple, they are selected/clicked on. The textual content shown in the interface is illustrative and does not resemble the potential of an LLM-based system in terms of quality and novelty.

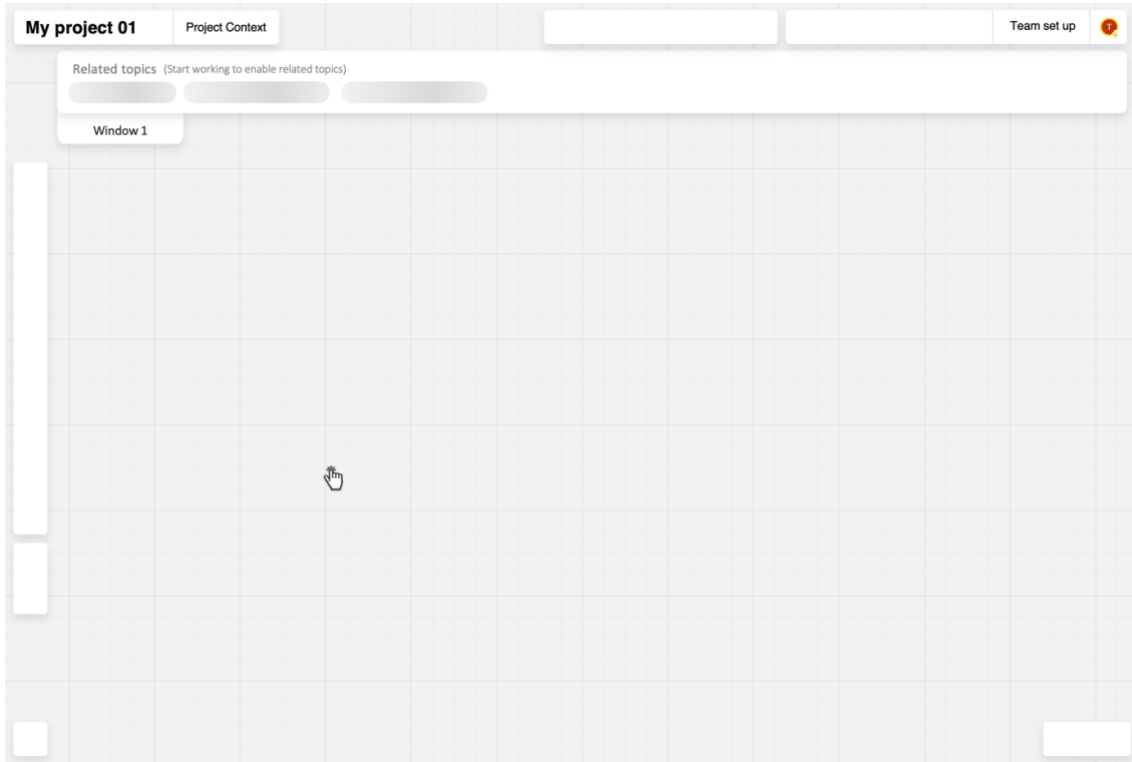


Figure 40: Digital working canvas

### 8.3.1 Strategy layer implementation

The first step is to provide the system with information about the project context, including the dimensions of project characteristics, involved stakeholders, timeline, and the designer's profile (figure 41). The 'project success parameters' and 'stakeholder success parameters' are regarded as the learning and design goals of the designer. The 'personal profile' further details the expertise of the designer and the *attributes* as defined in the 'human-AI collaborative design-space exploration framework' (Viros-i-Martin & Selva, 2021a). This information will serve as reference material for the personas and other system functionalities. For example, the AS' *information organization* and *perspective* interaction views build on the dimensions introduced in this *project context* window. This will be detailed further down this section.

My project 01 Project Context Team set up

Related topics (Start working to enable related topics)

Window 1

**Project characteristics**

Describe the overall design brief

Drag & drop project and/or design materials

Project success parameters

**Stakeholders**

Name

Main concern

Success factors

+

**Timeline**

From To

Date Date

**Milestones**

Title/Deliverable Date

+

**Your personal profile**

Describe your areas of expertise

Describe your perspective on this project.

Include in this descriptions your:

- Expectations & Preconceptions
- Attitude towards risk taking
- Personality characteristics

Figure 41: Project context window

### 8.3.2 Tactics layer implementation

The second step is to configure the personas (figure 42). Personas can be specifically created for the project or can be carried over from prior projects. The configuration is done by describing the purpose of the persona and providing some interaction characteristics. The content displayed next to the sliders is based on the ‘human-AI collaborative design-space exploration framework’ (Viros-i-Martin & Selva, 2021a) and design process framework. For example, the first slider establishes whether the persona augments in a top-down or bottom-up manner by defining whether its focus is on generating ideas or providing information. The slider ranging from ‘factual’ to ‘conceptual’ impacts the novelty of the persona’s propositions, as it defines the level of ‘hallucination’ that is allowed. In technical terms, this is the *temperature* of the LLM.

Apart from their own characteristics, the personas are provided with the project context including the design and learning goals. The result is that the personas are grounded in the same context as the designer, facilitating the alignment on collaboration goals.

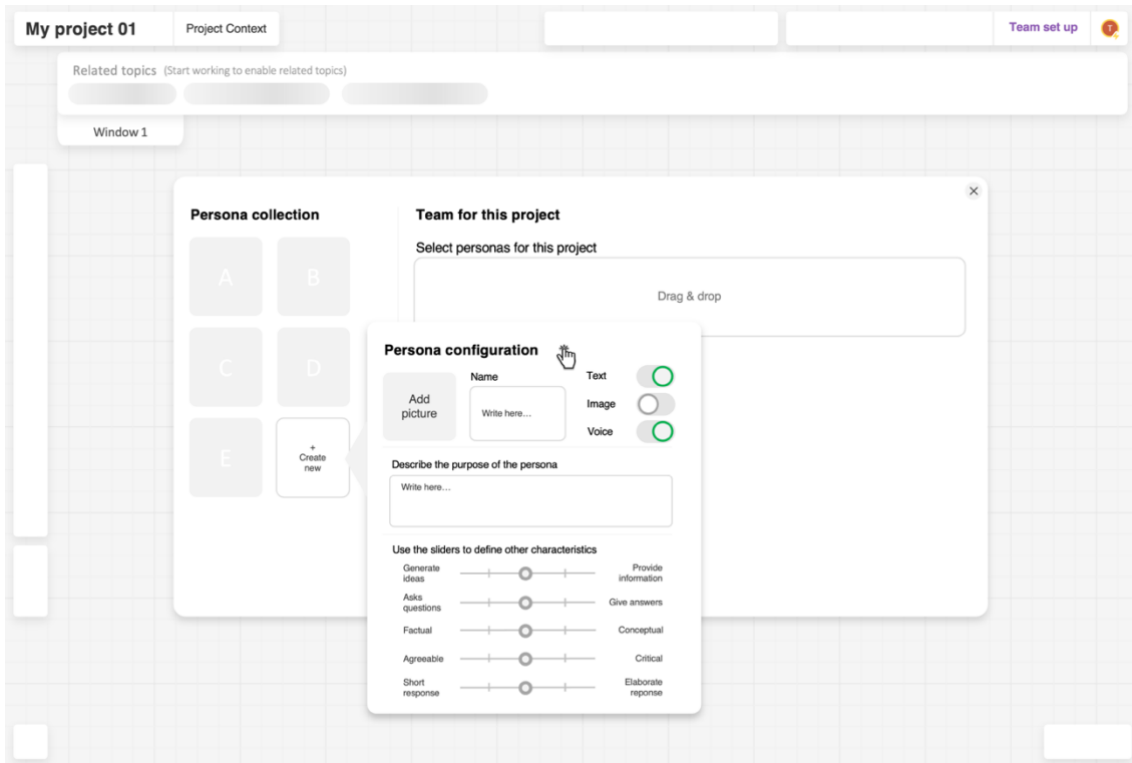


Figure 42: Persona configuration window

The main interactional interface component is the small selection menu that emerges when one right clicks in the working canvas (figure 43). The menu displays a couple of options. From top to bottom, the panel shows

- the different personas that one can converse with. The type of persona determines the form of augmentation. The interaction with personas is a form of representation action, as the interface projects information onto the canvas
- the 'discussion' option that allows for the combination of different forms of augmentation to address the limitations of top-down augmentation and allow for the combination of top-down and bottom-up based personas.
- the 'organize' option, which is the materialization of the *information organization interaction view*. It allows the user to switch between different views on the generated content in the working canvas. The re-organized presentation of design content is presented in a new tab that emerges under the top-bar once the *organize* option is used.
- the 'perspectives' option, which is the embodiment of the *perspective interaction view*. Here, the AS provides the user with feedback from a specific perspective on its generated content that is visible on the working canvas.

Each option will be discussed and visualized in more detail below.

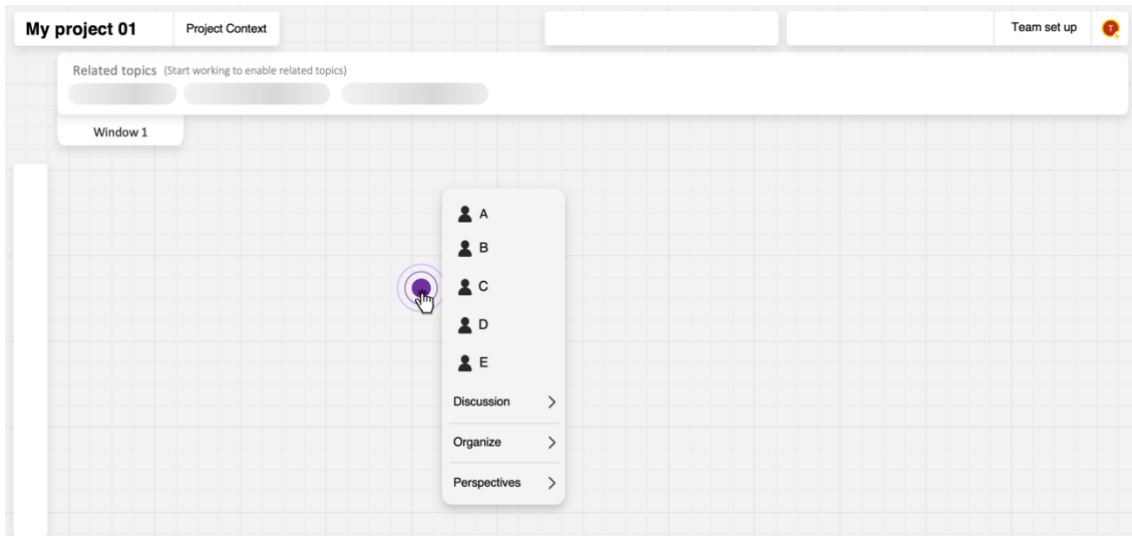


Figure 43: Main interactional interface component

### 8.3.3 Execution layer implementation

#### Personas

As mentioned in section 6.7.1, personas are the embedding of the six forms of augmentation. There are two ways in which the interface uses personas.

- First, there is the conversational use of personas on the digital canvas for the top-down and bottom-up types of augmentation.
- Second, the directional augmentation type is represented by the related topics as proposed in the top bar.

With regards to the first way in which personas are integrated in the interface, table 21 visualizes the interface progression after having selected a persona. The fourth column shows how the conversation, once being deselected, gets compressed into a smaller block that represents the content via a single title. Given the configuration of Persona A as visualized in figure 44, the interface thus represents a form of bottom-up quality augmentation provided by information and representation action.

Type window after selecting a persona

Typ, speak or share other content

The persona generating a response

What types are materials can be used for a window frame?

Persona A

The persona's responses

What types are materials can be used for a window frame?

Persona A

Window frames can be made from wood, vinyl, aluminum, fiberglass, composite, steel, or clad wood, each offering different benefits in terms of insulation, durability, maintenance, and cost.

Rank the materials mentioned from cheap to expensive

Persona A

Vinyl

Aluminum

Fiberglass

Composite

Wood

Clad wood

Steel

The conversation block once it is not selected.

Window frame materials x price

Material	Price
Vinyl	€
Aluminum	€
Fiberglass	€€
Composite	€€
Wood	€€€
Clad wood	€€€
Steel	€€€€

Persona A

Persona A

Table 21: Persona interaction windows

Persona configuration

Add picture

Name

Persona A

Text

☒

Image

☐

Voice

☒

Describe the purpose of the persona

The purpose of persona A is to be an analyst

Use the sliders to define other characteristics

Generate ideas

Provide information

Asks questions

Give answers

Factual

Conceptual

Agreeable

Critical

Short response

Elaborate response

Figure 44: Persona A configuration

Table 22A shows how such a block can be used as context for a conversation with another persona (left column). By drawing an arrow from one block into the empty space, the menu shown in figure 43 emerges again, allowing the user to choose the next move. Moreover, multiple blocks can be selected by a *context frame* (table 22B) to form the context for yet another persona (right column). As such, the *context frame* defines the tokens that become part of the context window of the next persona.

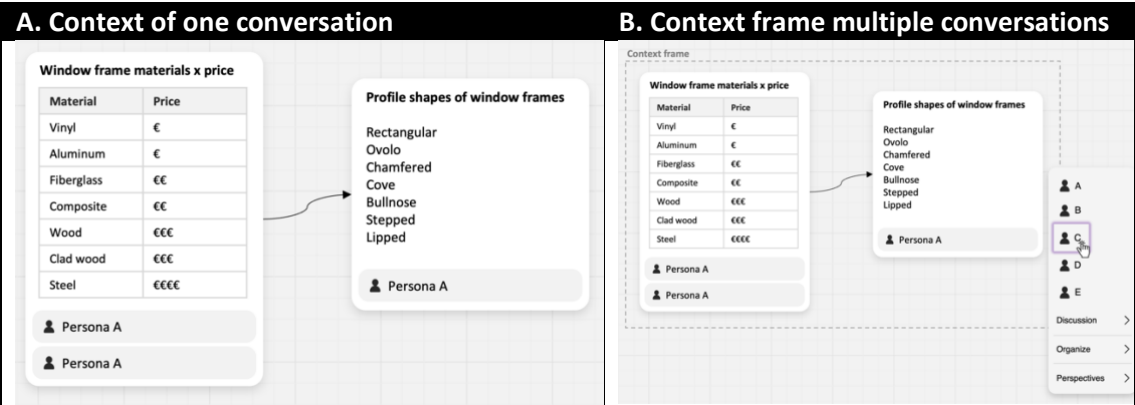


Table 22: Context window selection

With regards to the second way in which personas are integrated in the interface, figure 45 shows how, depending on the context, the system proposes related topics in the top bar. These topics are proposed by reflecting on the content that is presented on the working canvas relative to the description of the project brief and (stakeholder) success factors. As such, the top bar represents a form of *directional* augmentation that aims to inform the next action, and as such mental iteration, of the designer.

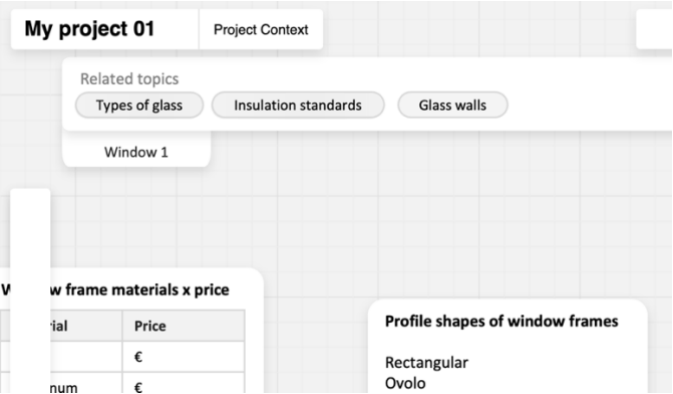


Figure 45: Related topics

### Discussion

The discussion option enables the designer to combine a selection of personas into a string of response-input pairs. Both the order of the personas as well as the conversation rounds can be configured (table 23A). First, the designer provides an opening question. Then the first persona responds to the question, where the response is the input for the second persona. This mechanism repeats when more personas are involved. When the final persona has generated its response, it becomes the input for the first persona again. This loop will be repeated for the configured amount of conversation loops. When the conversation rounds have been finished, the final answer that summarizes the 'discussion' is generated (table 23B).



This feature builds on two prompt-related insights from section 7.3. First, it builds on the idea that prompts can alter the characteristics of the output. This mechanism is embedded in the configuration of the personas, and as such enables the system to generate content triggered by various domains of semantic and syntactical features. Second, the hypothesis that iterative generation can improve the quality and novelty of the output is embraced by enabling the string of personas to form a loop. For example, given the assumption that persona A is configured such that it resembles an *analyst* who provides information, persona C is a free-thinking co-ideator that generates ideas, and persona B is a domain specific expert who critically evaluates the discussion, the flow of generated content follows a path similar to how the design process framework models a mental iteration.

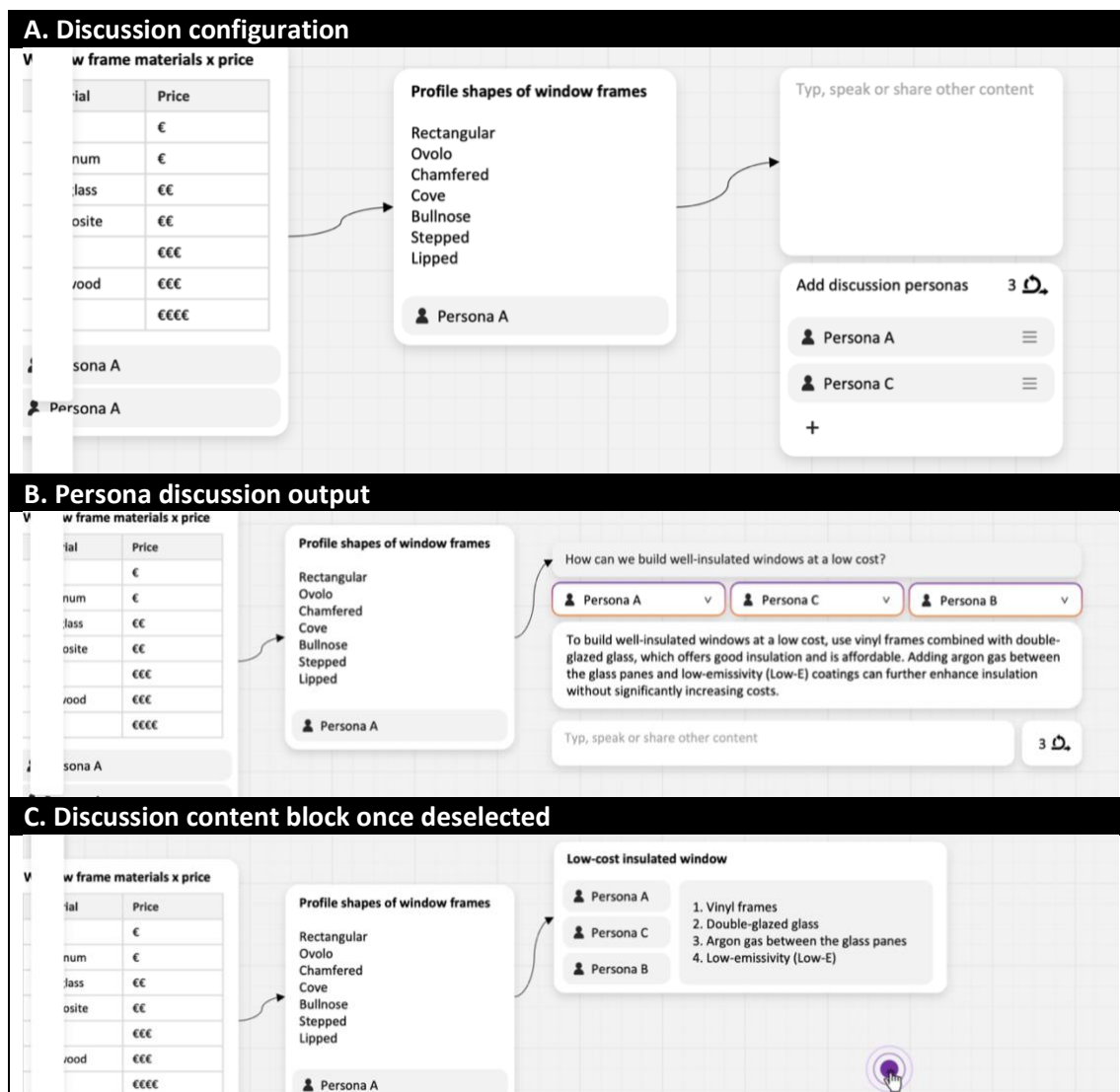


Table 23: Discussion interface progression

Organize

The *organize* option, supports the designer in visually structuring the content on the board. By drawing a *context frame* around the content one would like to organize, one can select a parameter that is used to structure that content (table 24A). For example, given the parameter ‘personas’ one can choose to organize all content per *persona* to identify what types of augmentation has been used most, and how the design process can benefit from an increased use of other types of augmentation or perspectives (table 24A). Every form of organization gets its own ‘tab’ (organize: Personas) underneath the top bar (table 24B).

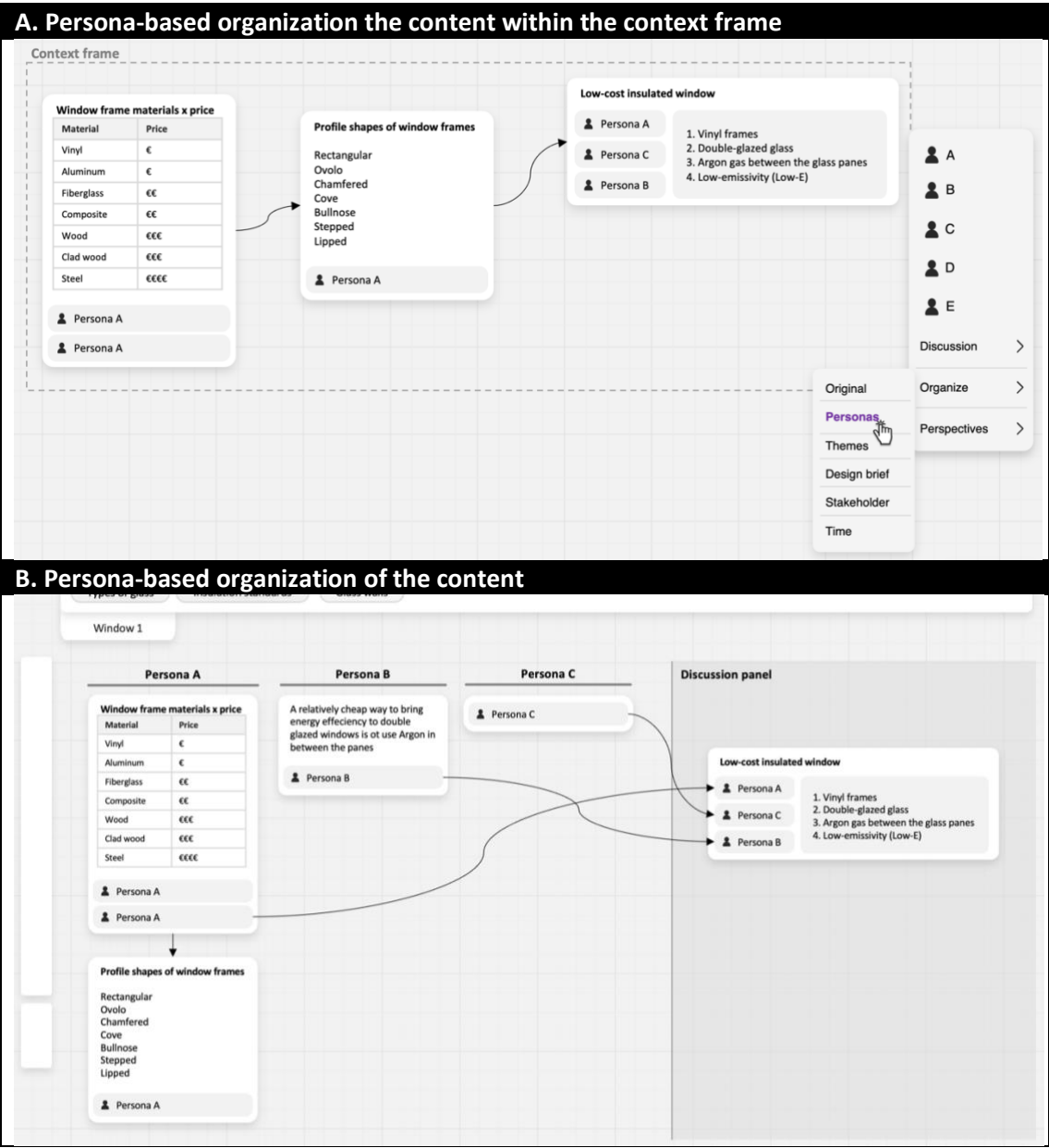


Table 24: Content organization

Another form of organization is the design brief and its success factors. Table 25 gives an example of how such organization can work. Here the project success factors, as configured in the *project context* (figure 41), are connected to already generated content. Per project success factor, suggestions on missing content can be made by personas. These suggestions can be added to working canvas (table 25B).

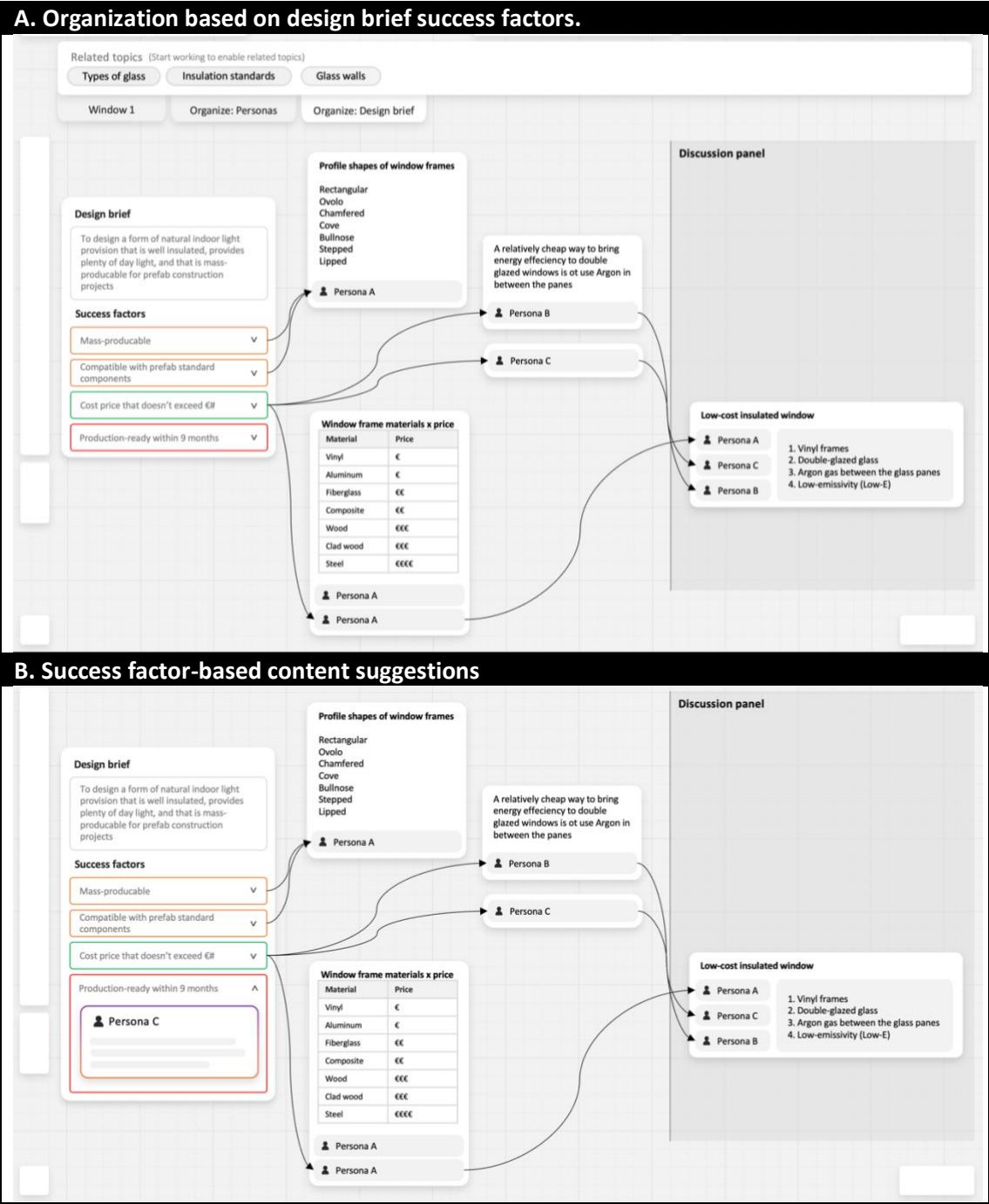


Table 25: Content organization

Perspectives

The *perspectives* option represents a new layer on top of the *information organization* interview view. Similar to how *knowledge sharing* action is a descriptive variant of the reasoning behind *information* action, the *perspective* option scans over all the content to build a line of reasoning towards what elements can still be improved or are missing. Based on the stakeholder profiles (table 26A), as configured in the *project context* (figure 41), the system can provide feedback or critical questions on the content as projected on the working canvas and embed that on a content-block basis (table 26B).

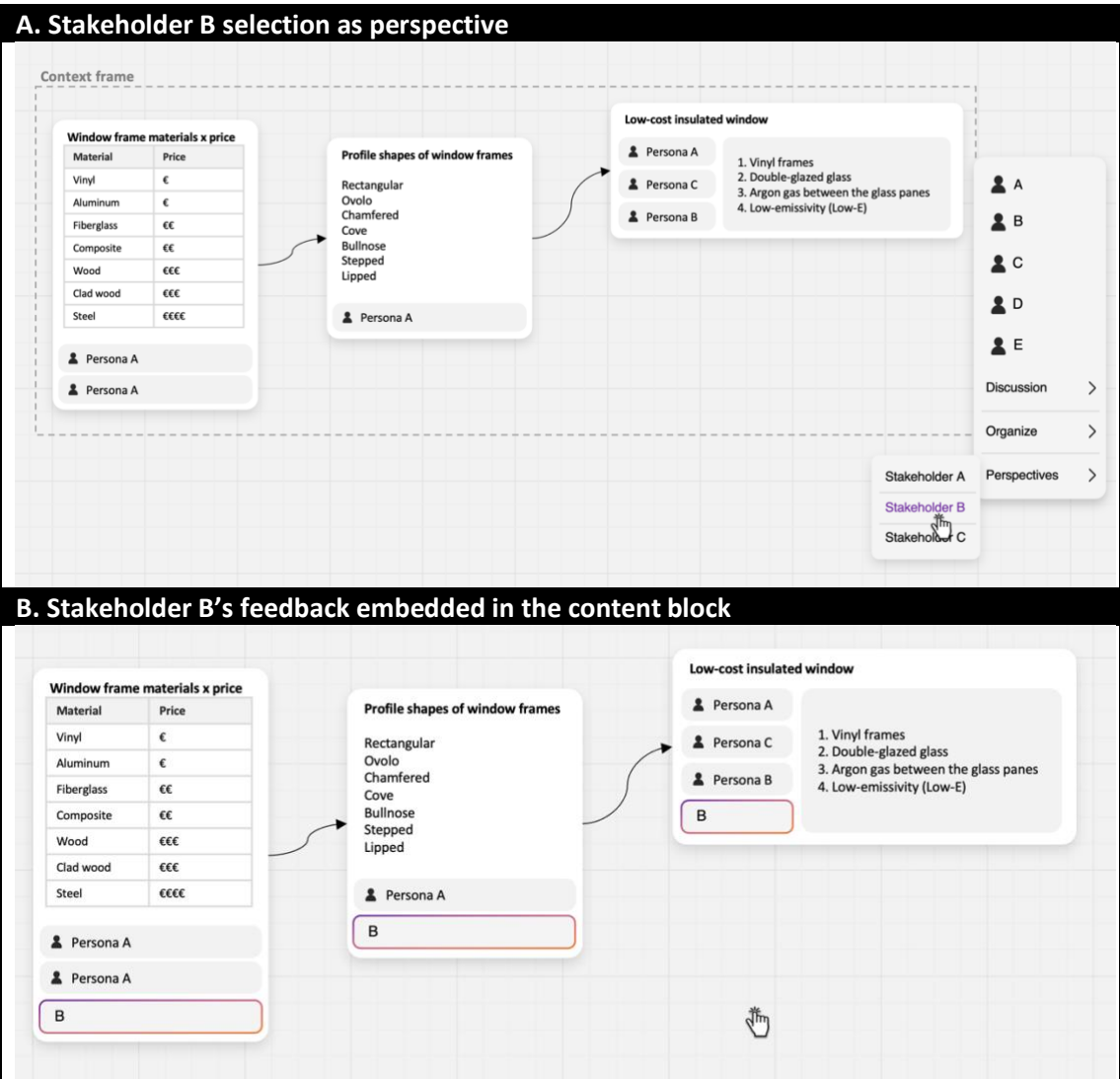


Table 26: Stakeholder perspectives

## 9. Socio-economic implications

---

Based on the presented augmentation system interface, one can imagine various potential ways in which such a system can be used in practice. In this section, potential pathways are introduced followed by a discussion on the human role relative to that of the augmentation system, the implications for organizations, and their feasibility and desirability. As such this section attends to the sixth research question:

*What are the socio-economic implications of LLM-based augmentation systems?*

---

### 9.1 Persona-flow patterns as an organizational operating model

The augmentation system (AS) interface as shown in the previous section, allows the user to work with a set of self-created personas that act as a virtual team that augments the user of the AS. Besides organizing the personas in a certain custom way for every project, one can also imagine a scenario where either the AS or the human individual identifies patterns in persona flows over various projects. For example, the observation may be that at the start of a project persona A is used a couple of times, to then use the generated content to feed it to persona C, and to finally feed all content to persona B who wraps up the information into a concept proposition. This insight can then be used to use certain persona-flow patterns. These ‘standard’ patterns can then be employed and slightly customized to work not only for their original project, but to also work in new situations. Often, such structures are called agent frameworks (as in (Liu et al., 2023) and (Wu et al., 2023)), where within a certain confined space of objectives, agents are able to autonomously collaborate and execute.

A clear similarity can be observed between the augmentation, or even the automation, of the individual through these persona-flow patterns and the individual as augmented by the structure of an organization. Augmentation at the scale of organizations is realized by organizing human expertise and knowledge in a structure such that the objective of the organization, often realized by the board or (top) management, is best materialized. In the case of a designer being augmented by an LLM-based system, a similar relation exists between the objective of the individual and the structure of information, embodied in various agents or personas, that surrounds it. This observation points at two questions. What role do we assign to ourselves relative to the role of GenAI-based systems? What will the future of organizations look like? The next section covers the first question, the section after explores the second question.

## 9.2 The role of humans and augmentation systems

One way to approach the first question – what role do we assign to ourselves relative to the role of GenAI-based systems? – is to look at the layers of behavior as defined by Cash & Kreye (2017) and implemented in the design process framework in figure 28. These three layers do not only describe the *actions* that are executed, but also the action-steering *task* and *activity* layers. The question then becomes what layers we humans want to control. Building on the concept of hybrid intelligence (Dellermann, Ebel, et al., 2019) it can be argued that, as humans, it is desirable to be in control over the goal-oriented and motivation-driven *activity* layers. The *tasks* layer, containing the consecutive tasks that attend to the various simultaneously active *activity* layers, can then be a hybrid playing ground between human and AS. Finally, depending on the type of activity, one may choose to (A) perform the *actions* layer oneself, (B) fully ‘outsource’ the *actions* layer to the AS with certain governance practices surrounding the execution, or (C) to perform a sequential form of interworking with the AS (figure 46). It is important to note that in scenario B, governance is still highly relevant as illustrated by the AI paperclip problem<sup>20</sup>. Similarly, the orthogonality and convergence theses argue that benevolence is not to be assumed part of the intelligence of a computational agent (Bostrom, 2012) .

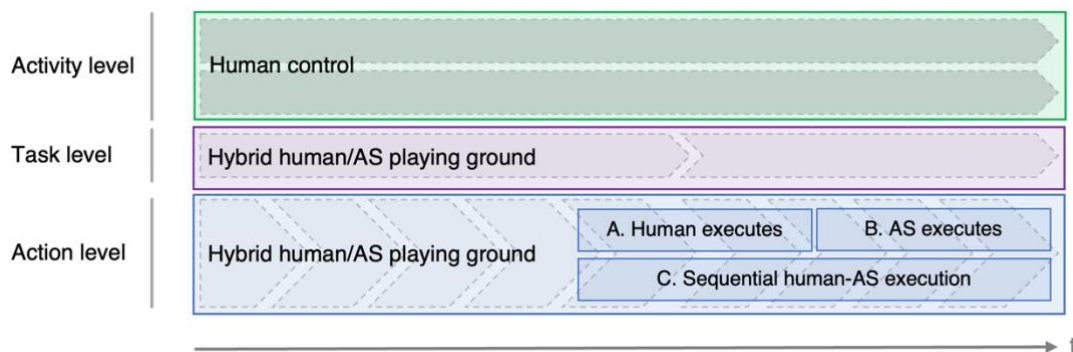


Figure 46: Layers of behavior and human/AS control

## 9.3 Implications for organizational structures and data

The next question is what the future of organizations can look like given these different forms of augmentation. The implications of these forms of augmentation depend on the type of organization and its objectives.

In a large hierarchical organization (e.g., multinationals) with a base of, for example, ‘analyst’ functions, there are two hypothetical options. First, the persona-flows (or an agent framework) may replace a set of analysts, with only a (project) management function providing objectives (*activity* layer) to the augmentation system, for it to then

<sup>20</sup> <https://medium.com/@jeffreydutton/the-ai-paperclip-problem-explained-233e7e57e4e3>

semi-autonomously follow certain paths of analysis and present the outcomes (*tasks* and *action* layers). This corresponds to an AS that executes *actions* autonomously and mostly controls the *task* layer as well. The second scenario is one where the number of analysts remains the same or is reduced slightly to either increase the organization's output or become more efficient respectively. This scenario represents an AS that collaborates with humans in the definition of *tasks*, and augments or semi-automates the execution of *actions*. From these two hypothetical scenarios, the first one represents a form of semi-automation, whereas the second one can be regarded as augmentation.

In smaller organization with broader role definitions, a similar trajectory may unfold although there are less clearly defined boundaries as to which roles represent what specific functions. Like the large hierarchical organization, the choice for automation versus augmentation depends on the objective of the organization. The scenarios are summarized in figure 47, showing the two ways in which a GenAI-based augmentation system may change the way in which organizations work, including the underlying objectives and requirements.

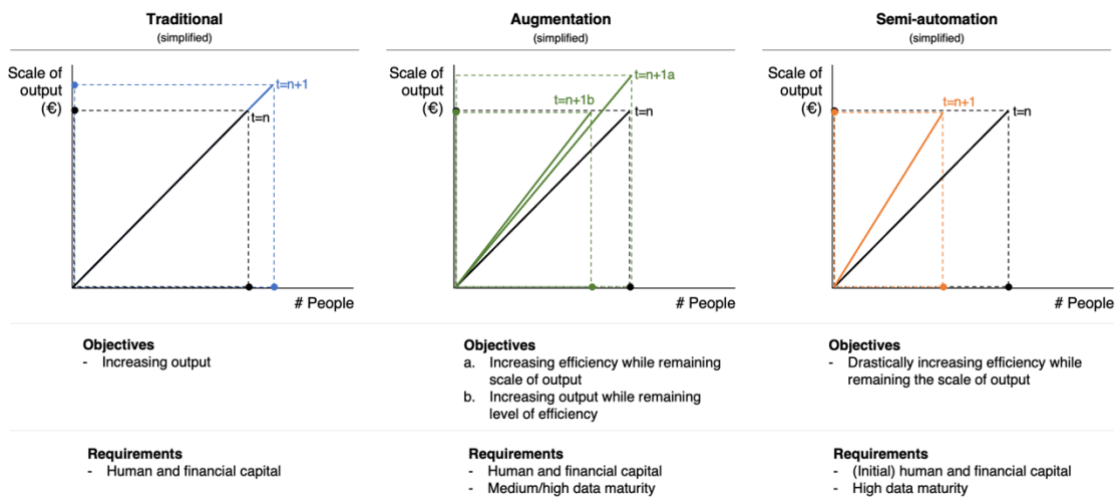


Figure 47: Organizational implications, objectives, and requirements

#### 9.4 Feasibility: data maturity

Reflecting on the topic of feasibility based on my experience as an intern within a professional services company in the business-IT domain, the feasibility of the AS applications provided in this section mostly depends on the data maturity of an organization (apart from aspects such as the purpose and values of the organization). Related topics include, among others, setting up the right data collection points, structuring the data, storing data in a central and secure location, and having the data governance practices in place that ensure data privacy and further regulatory compliance in access management for both internal and external stakeholders (including customers).



For example, if a designer at company A wants to communicate with a digital ‘archetype’ or ‘specific’ customer of company A to better understand a certain need, the LLM-based system needs to have access to documents with recordings of customer feedback or needs. This document has to be recorded in a structured format and saved in a central place, such that either functions from all over the organizations or a set of *augmentation personas* can access it. Finally, the filing of customers’ data needs to be compliant with any privacy related policies. This may introduce certain limitations on the way in which information can be represented by *augmentation personas* and shared within the organization.

The timeline for implementing such AS applications often requires large-scale transformations, which can take several years before large multinationals, for instance, can fully adapt their organizations as described in this subsection. In the shorter term (<1 year), specific smaller parts of the organization may be able to experiment with and build towards a first version of augmentation at the level of semi-automating small parts of their workflow. Instead of large incumbents, the short-term realization of augmented and/or semi-automated organizations may come from smaller and/or new companies that can build their organization from scratch based on the principles of augmentation and automation as described above<sup>21</sup>.

## 9.5 Desirability

Although the technological advancements in the domain of GenAI provide increasingly feasible and viable options for both the augmentation and automation of some parts of the human cognition, the question remains if these developments are desirable and thus if they should be adopted. If this question would be proposed to an imaginative person who represents the free-market capital system, the answer would be that every opportunity that enables organizations to become more efficient should be embraced as, ultimately, the consumer will benefit because of the lower prices. However, one could argue that introducing measures that replace human cognitive action requires reflection on a scale larger than the narrow scope of capitalism that essentially reduces the complex construct of the ‘human experience’ to ‘being a consumer’ and ‘progress’ to ‘efficiency gains’, while in the process of doing so externalizing negative consequences such as Co2 emissions. Below a set of expectations, defined at various timescales, is listed given the scenario of an AS that either automates or augments an organization.

---

<sup>21</sup> This prediction is based on several podcasts from researchers or industry professionals that I’ve listened to over the years. A couple of recent examples include: All in Podcast ([Episode 28-09-2024](#)), Lex Fridman podcast ([Episode: 19-06-2024](#)) ([Episode 18-03-2024](#)), Dwarkesh Patel ([Episode 19-05-2024](#)) ([Episode 28-03-2024](#)) ([Episode 04-06-2024](#))



In the short term, the main societal impact can be a shift in the industries' human capital demand, leading to the revaluation and potentially loss of certain jobs. It is not yet clear what type of new jobs might arise<sup>22</sup>. In the medium term, a discontinuity might emerge between the current education system and the demand from job markets. This could require a transformation of the way in which we look at the basic knowledge and skills an individual should have to become an effective economic contributor. Finally, in the long term, the organization of human cognition around a central pool of data, or the complete automation of it, might increase the risk of a global monoculture and the loss of diversity in perspective<sup>23</sup>. If everyone has access to the same information, the possibility that the everyone constructs the same (logical) accounts of reasoning to fill a gap in their understanding increases. Compared to a scenario in which different groups have a different set of information as a starting point for the pursuit of further knowledge exploration, this might lead to less diverse outcomes. Further contemplation on this topic is required to better understand the dynamics and hence the consequences of such a scenario.

Although the previous paragraph presents hypothetical scenarios that are open to different interpretations, it points at the level of societal disruption and system change that is related to questions about augmentation and automation. The real question that should be asked is not whether to go for augmentation or automation, but how much societal disruption we are willing to tolerate in the name of economic efficiency gains. What level of potential societal disruption and system-change is desirable at what point in time and how should the potential economic efficiency gains stack up against it?

#### **Takeaway box 16**

Besides using personas-flows for augmentation, they can also be used to automate the work of a(n) (set of) individual(s). This presents an interesting space for debate on the role of humans relative to GenAI, and relative to the functions within organizations. On a high level, there are two ways in which AI can be employed leading to either augmentation or automation. Depending on organizational objectives (efficiency, output growth) both scenarios may occur. Each scenario brings with it feasibility considerations about data maturity, as well as considerations on desirability. The latter type of considerations includes implications such as a potential shift in human capital demands, a disconnection between the education systems and labor markets, and (proposed tentatively) an increasingly monotone global information system leading to less diversity in human creations.

---

<sup>22</sup> As reported in a study from Ernst & Young (EY): [Link](#) to report

<sup>23</sup> A statement based on a discussion during the focus group

## 10. Discussion

---

The central research question of this thesis is how to augment the design process using LLMs, driven by the pursuit to contribute to the conceptual understanding of GenAI- and specifically LLM-based design process augmentation. To answer this question, this thesis has generated a framework of the design process as a platform for design augmenting practices, a framework for structuring various types of design augmentation practices, an assessment of LLM design process augmentation capabilities, and an interface for designer-LLM co-design. The following sections address both the theoretical contributions of this thesis to the existing theoretical landscape and its limitations. Moreover, it provides some final reflections on the implications of this thesis' findings for the design practice in general and presents some pointers for future research avenues.

---

### 10.1 The research approach: putting the design process front and center

In general, the approach of this thesis has been the differentiating factor compared to other accounts of research. Stated in an abstract way, this thesis has flipped the perspective on this technology-based way of augmenting a cognition-heavy human process by not taking the technology and the interactions that it affords as a starting point, but by looking at the ways in which the human cognition in its process of design could be enhanced. Thoring et al. (2023) took a similar approach. However, their research was inherently limited by the depth of analysis and formulation of the central building block: the design process. In contrast to other research that has been either empirical or conceptual in nature and that used existing human-AI interaction patterns and technology-based characteristics to identify how or to come up with ways in which the design process could be augmented (Berni et al., 2024; Chen et al., 2024; Dellermann, Calma, et al., 2019; Fabri et al., 2023; Lawton, Grace, et al., 2023; Lawton, Ibarrola, et al., 2023; Li & Li, 2024; Lin et al., 2023; Y. Liu et al., 2024; Schwartz et al., 2023; Van Zoelen et al., 2021; Viros-i-Martin & Selva, 2021b), this thesis has formulated a new and unified perspective on the design process that attends to most of the person-related characteristics as described in a human-AI co-design framework (Viros-i-Martin & Selva, 2021) as a basis for reasoning about ways in which it could be augmented. The benefit of this approach is that it isn't constrained by either the sophistication of existing human-AI interaction literature or the capabilities of existing technologies. As such, this thesis has created a more fundamental categorization for forms of design process

augmentation in which the technology-based projections of design augmentations can be further embedded.

## 10.2 Theoretical contributions & limitations

### 10.2.1 Design theory

#### *Contributions*

With regards to the first two gaps in research – the lack of descriptive power of individual design theories and the lack of connection between design theories hindering the overall connection between the design process and the domain of human-AI collaboration – a unified design process framework has been conceived that attends to most of the person-related factors of the collaborative human-AI design space exploration framework from Viros-i-Martin and Selva (2021). The implications of the unified design process framework are described below.

First, the factor that hinders the progression of design theory formulation – the lack of a shared vocabulary and ontology (Viros-i-Martin & Selva, 2021) – has been resolved. The main design theories, such as problem-solution co-evolution (Cash et al., 2023; Crilly, 2021b, 2021a; Maher & Poon, 1996), the situated function-behavior-structure ontology and framework (Gero, 1990; Gero & Kannengiesser, 2004), and the concept-knowledge theory (Hatchuel & Weil, 2002), have been unified by grounding them in the theory of *schemata* as knowledge representations (Rumelhart & Ortony, 2017), visualized by active structural network representations. The result is that a basis has been conceived for further theoretical integration by providing a clear view on the similarities and differences in the individual theories' relative descriptions of design.

Second, by relating the design theories mentioned above, this thesis has proposed an alternative definition of co-evolution in design to that of problem-solution co-evolution. By reintegrating the space of *knowledge* into the equation – a space that got lost in the translation from the original biological origin of the theory itself (Crilly, 2021b) – the co-evolving spaces of *knowledge*, *memory*, and *concepts* have been constructed. Interestingly, the spaces of problems and solutions can be defined as a derivative of these more fundamental spaces. This contribution is in line with Crilly (2021b) who described how future versions of the co-evolution model in design could look into the addition of more spaces that can co-evolve, where such spaces could also be 'nested' into each other at different scales. The design process framework presents an example of 'nested' spaces, with the spaces of *proposition* and *understanding* being nested in the space of concepts.

Third, the framework not only unites existing theories, but also expands them into the domain of collaboration, as such meeting the demand for conceptual expansion towards the collaborative domain (Cash et al., 2023; Crilly, 2021b). Since the unified

design process framework is connected to literature about human-AI collaboration, design theory now has the potential to expand beyond the scope of understanding the design process at the scale of an individual towards understanding it at the scale of two or more interacting agents.

#### *Limitations & future research*

One of the limitations of the unified design process framework is its lack of integration of person-related attributes, although such factors are described as important in literature about the progression of human-human (Fyhn et al., 2023) or human-AI (Viros-i-Martin & Selva, 2021) collaborative processes. Further conceptual integration is required in this domain. Future (empirical) research aiming to design research at the intersection of individual-specific traits and design process progression should either control for these variables to minimize their effect or perform tests to better understand their relationship relative to the process of design. Furthermore, although this thesis argues that the design process can be described as a process of propositional learning, the way in which the interpretations of an individual are translated into new knowledge remains largely undefined. As also identified by Gonçalves<sup>24</sup>, it could be useful to connect the design process framework to Bloom's taxonomy of learning (Adams, 2015). Another limitation is the fact that there is not yet a clear connection between perceived uncertainty, being grounded in one's meta-cognitive processes, and the tacit formulation of goals. As such, contrary to design process progression, the direction of the design process cannot be described as derived of one's cognitive activity. Future research can more closely look into this area to not only describe the driver of the design process but also the factor that establishes its direction in terms of an individual's meta-cognitive processes.

Moreover, moving to the collaborative domain, there is not yet an account of the relation between an individual's collaboration goals (as defined in the framework from Viros-i-Martin and Selva (2021)), emergent states, as described in literature about teamwork effectiveness and progression (Fyhn et al., 2023; Grossman et al., 2017; Marks et al., 2001), and the design process. Emergent states include factors such as collective efficacy, potency, cohesion, and situational awareness, and are seen as both 'products' (or outcomes) of team processes and inputs to subsequent processes and outcomes (Marks et al., 2001). As such, the human-AI co-design framework is still limited in its ability to describe the effect of one's actions on the state of the other actors. However, due to the structure of mental iterations, the integration of epistemic perceived uncertainty (Cash et al., 2023; Cash & Kreye, 2017), and the connection to the Uncertainty Driven Action (UDA) model of design (Cash & Kreye, 2017), there are various entry points for further conceptual expansion in this direction.

---

<sup>24</sup> An insight that was presented during the focus group

In general, the framework remains a theory that requires empirical evidence as proof of validity. As such, future research should focus on designing empirical experiments to test the underlying premises of this design process framework to further enhance its validity.

### **10.2.2 LLM-based design process augmentation**

#### *Contributions*

With regards to the augmentation of the design process, this thesis has proposed a framework that describes the six forms of design process augmentation including the three different ways in which an Augmentation System (AS) can execute those. Based on this framework, an investigation into the design process augmentation, and thus creative, capabilities of LLMs has been executed. Creativity frameworks were used as an instrument for the structured construction of a perspective on the creativity of LLMs. The findings of this thesis contribute to areas defined as critical for future research in the area of AI-augmented design. First, Thoring et al. (2023) present a research agenda containing ten questions about the integration of GenAI in the design process. Nine out of the ten questions rest on a deeper question about how to create a GenAI-based system that both aligns with the desired direction of the designer and generates novel and quality suggestions in an iterative manner. This thesis has addressed both points. First, the design process framework, including its three layers of behavior, describes both the driver of the design process – meta-cognitive perceived uncertainty – as well as the factors that influence the direction of the process – learning, design, and collaboration goals. As such, it provides a theoretical basis for reasoning about how to align designer and augmentation system. Moreover, by providing a view on the creative capabilities of LLMs, this thesis has also addressed the call for a better perspective on GenAI’s capabilities to reinforce human creativity. Finally, with regards to the exploration of both automation and augmentation paths, this thesis has created the basis for Thoring et al.’s (2023) call for future work that focusses on *“the relation of Generative AI to the design process”* (p. 3352).

Second, Berni et al. (2024) state that *“the presumed AI capabilities of disclosing links between stimulation means and designers’ actions [...] have not been exploited and turned into desired design applications”* (p. 1944). By establishing a clear and more complete overview of person-related characteristics in the design process framework (section 5), the definition of augmentation forms (section 6), interaction functionalities as defined in both AS architecture (section 8.2), and the proposed interface (section 8.3), this thesis sheds a light on what such an application could look like and how it could work on top of LLMs. Moreover, adding up all these factors (the design process framework, creativity assessment of LLMs, and the augmentation system architecture and interface) this thesis has grounded a predictive view on the ideal augmentation system from Liu et al. (2024) in the relevant theory and literature. In the words of Liu et al. (2024): *“an ideal [AI-based system] design would likely involve a dynamic adjustment*

*of AI's role based on the user's expertise, background, and desire for control, allowing for both an engaging co-creation experience and innovative outcomes" (section 8.1.2).*

Besides calling for clarification on what is theoretically possible, this thesis also aligns with the ethical considerations as embedded in the framework of AI-support for stimulation in design ideation from Berni et al. (2024). For example, the augmentation system (AS) as proposed in section eight ensures that the designer is the central decision-maker, that the AS does not replace the designer, and that the AS does not nudge the designer in a specific, limiting direction, leading to early design fixation.

Compared to two existing LLM-based systems, the Supermind Ideator (Rick et al., 2023) and Allixr<sup>25</sup>, the augmentation system (AS) interface proposed in this thesis is more theoretically grounded in design theory literature. In line with the research objective<sup>26</sup>, the proposed AS interface is not designed based on the taxonomy of interface features. It takes the design process as central element in conceiving an interface that augments it. Moreover, the proposed AS interface is better integrated into the workflow of a designer. By realizing what Davis et al. (2015) call participatory sense-making by working with *personas* and different *interaction views*, the proposed AS interface embeds the most fundamental building blocks of the design process in a set of interaction components that represent the basis for the collaborative process between designer and LLM-based AS system. In contrast, the Supermind Ideator (Rick et al., 2023) allows the user to either "*explore problem*" or "*explore solution*" based on a given goal (figure 48). It then generates in one go a set of aspects and related topics that it considers important for figuring out how to realize that goal. One could argue that when the augmentation system generates all its content in one go, the integration between a designer's thoughts and actions is not taken into account. Besides, Allixr's interface presents a set of design widgets that each cover a specific aspect of the design process (figure 49). However, these design widgets do not cover all aspects that a designer mind encounter during his process. Moreover, by presenting these widgets as a separate category specifically focused on supporting the design process, the act of searching for information and organizing information is implicitly positioned as a different type of activity.

---

<sup>25</sup> [Link](#) to the corresponding master thesis from Ruben Dekker (2022)

<sup>26</sup> Recalling from the introduction: the research objective is to contribute to the conceptual understanding of design process augmentation by defining 1) how humans design, 2) what interaction can augment that design process, and 3) what augmentation system interface can realize the desired forms of augmentation.

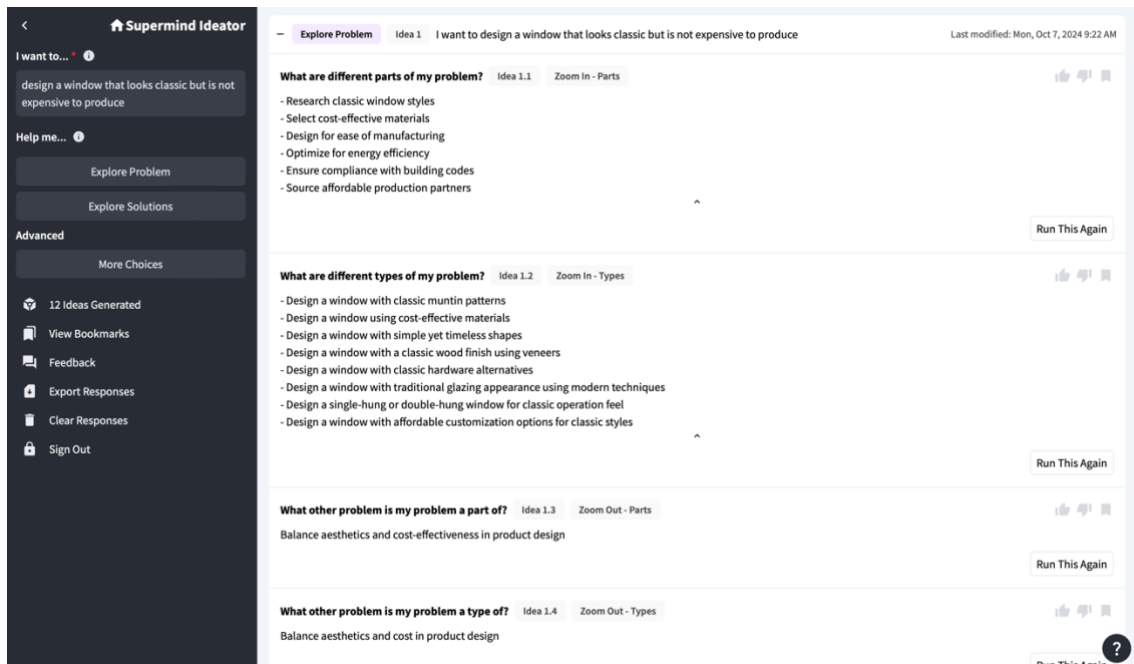


Figure 48: A screenshot from the Supermind Ideator interface (Rick et al., 2023)

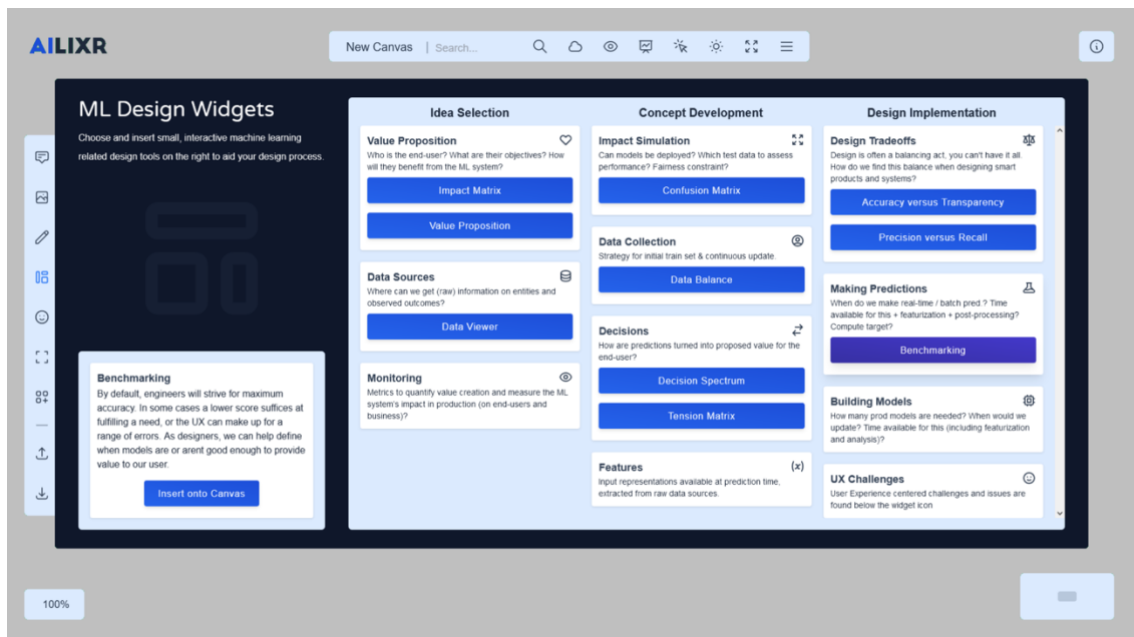


Figure 49: Screen shot from Ailixr's design widgets ([Link to thesis](#))

### Limitations & future research

There are some limitations with regards to the contributions in this domain. First, the contributions are theoretical in nature and, due to the lack of a functional interface, lack insights based on empirical user testing. Second, as mentioned in the introduction, the aim is to contribute to the understanding of how forms of GenAI can augment the design process. This thesis has only focused on LLMs, and as such has not investigated the

integration of Generative Adversarial Networks (GANs) or other forms of transformer-based generative systems that can, for example, be used for generating images. Third, the proposed interface for the augmentation system is based on the screen-based monitors. Especially with the very recent advent of Augmented Reality (AR) glasses from Meta, the Meta Orion<sup>27</sup>, and existing VR products such as the Meta Quest 3s<sup>28</sup> and Apple Vision Pro<sup>29</sup> it is also possible to start experimenting with these new interface types. Fourth, with the reveal of OpenAI's o1 model<sup>30</sup>, the capabilities of LLMs have been extended beyond the capabilities discussed in this thesis. Although this might result in outdated findings in the investigation of LLM creativity, it does not change the fundamental propositions of this thesis in the domain of design theory and augmentation.

### **10.2.3 Socio-economic implications**

#### *Contributions*

Section nine highlights the practical ways in which augmentation systems might be deployed and their effect on the economic output of organizations, economic value of human labor, and further societal disruption. The layers of the human behavior framework act as a means to design Augmentation Systems (ASs) and understand the consequences of design choices. By reflecting on the different forms of human-AI interworking in the context of the economic input and output of organizations, a connection has thus been made between augmentation system design choices and potential socio-economic consequences, depending on the type of application and context.

#### *Limitations*

Although this is relevant material in discussions in different domains and at various scales, further research is required to detail the exact domains and the extend of impact within them. At the scale of academic research, the findings are relevant for the domains of economics and business/public administration. Moreover, research in the domains of cognitive psychology, anthropology, and history may want to investigate the long-term consequences of the organization of human intellect around a single (or small set of) centralized global information systems for the future development of human culture(s) and society(s). At the scale of countries, or continents, the findings may inform the debate about AI technology regulations, legislation surrounding the topic of human labor rights and wages, and the future of the education system.

---

<sup>27</sup> [Link to Meta Orion](#)

<sup>28</sup> [Link to Meta Quest 3s](#)

<sup>29</sup> [Link to Apple Vision Pro](#)

<sup>30</sup> [Link to OpenAI o1](#)



# 11. Conclusion

Current literature that aims to describe the use of (Gen)AI in the domain of design has been hindered by the lack of conceptual integration of design theories which in turn blurs the connection between descriptions of the design process and frameworks that describe the actor-related characteristics that influence and shape the process of human-AI interaction and collaboration. In line with the ethical consideration proposed by Berni et al. (2024), this thesis' objective is to contribute to the conceptual understanding of design process augmentation by defining 1) how humans design, 2) what types of interactions lead to the augmentation of the design process, and 3) how an LLM-based Augmentation System (AS) can realize the desired forms of augmentation. This thesis has shown how such LLM-based design process augmentation could look like by establishing 1) a time-based framework of a human-AS collaborative design process, 2) a perspective on the design process augmentation capabilities of LLMs, and 3) an augmentation system (AS) architecture and interface for the practical implementation of these theoretical considerations. Other contributions that have enabled these three main contributions include 1) the unification of existing design theories (problem-solution co-evolution, situated Function-Behavior-Structure (FBS) framework, Concept-Knowledge theory, and Uncertainty Driven Action (UDA) model) to attend to a wider range of person-related characteristics required for the description of a collaborative process of design, 2) the connection between human intellect augmentation and creativity literature, resulting in 3) the application of creativity literature in the domain of GenAI, and specifically LLMs, and 4) the use of the layers of behavior, as described in the UDA (Uncertainty Driven Action) model, as the basis for the augmentation system's architecture. The results of this thesis have created a strong fundament for further research into the direction of the initial objective, as the results have increased the conceptual power to describe and explain process, output, and other augmentation-related phenomena connected to the domain of human-LLM co-design. Finally, a reflection on technical feasibility and socio-economic desirability of different forms and applications of augmentation systems is provided. This leads to the conclusion that, instead of focusing on the question whether we want AI-based augmentation or automation, it is more relevant to focus on the question how much societal disruption we as humans would tolerate relative to the economic efficiency gains proposed by human intellect augmentation and/or automation. Future research in this domain should explore three key areas:

## *1. Expansion of design theory towards the collaborative domain*

With regards to the theoretical framework on collaborative design progression, the integration of person-related attributes remains underdeveloped. It is important to investigate how individual-specific traits influence the design process, particularly in the context of human-human and human-AI collaboration. Additionally, while this work has proposed the design process as a form of propositional learning, further study is needed

to understand how individual interpretations are transformed into new knowledge. Connecting this process to Bloom's taxonomy of learning could offer useful insights. Another critical area for investigation is the relationship between meta-cognitive processes, perceived uncertainty, and the formulation of design goals. Understanding this connection could clarify how cognitive activity shapes the direction of the design process. Moreover, in collaborative settings, future research should focus on the interplay between individual collaboration goals and emergent states—such as collective efficacy and cohesion—and how these factors influence design processes within teams. Besides, there is a need to further refine the human-AI co-design framework to better capture how the actions of one actor affect the state of others. Finally, apart from these theoretical concerns, the framework requires empirical evidence as proof of validity.

## *2. Empirical testing of GenAI-based interfaces*

With regards to the practical implementation of this conceptual framework in a GenAI-based augmentation system and interface, future research should focus on empirical user testing by using a functional interface to both validate the theoretical contributions as well as gather practical feedback. Additionally, further investigation is required to explore the capabilities of the new generation of LLMs, currently represented by OpenAI's o1, and to explore the possibilities presented by image-generation models (both transformer-based and Generative Adversarial Networks (GANs)). Finally, given the developments in the domain of new interface technologies, there is an opportunity to experiment with devices such as Augmented Reality (AR) glasses and Virtual Reality (VR) headsets to assess their potential of acting as useful interfaces for design process augmentation.

## *3. Understanding and dealing with socio-economic consequences*

Considering the socio-economic consequences of effective human-intellect and/or design process augmentation systems, it is important to better understand their effects in the short, medium, and long term. Future research should focus on ensuring both social and economic security of humans, by providing clear guidelines for regulations on the development and deployment of AI technologies (short term), human labor rights and wages (medium term), and the future of the education system (medium term). Research in the domains of cognitive psychology, anthropology, and history may want to investigate the long-term consequences of the organization of human intellect around a single (or small set of) centralized global information systems for the future development of human culture(s) and society(s).

Let us not waste our energy to naively pursue the creation of a set of entities that will replace us for the sake of short-term personal recognition and economic gains. Let us respect the capabilities of computers and complex algorithms but let us also celebrate the capacities that are unique to the human experience.

## 12. Reflecting on the (strategic) design practice

---

In this section, I sketch a more personal discussion of the content of this thesis. More specifically, this section aims to reflect on the theoretical contemplations on design theory and the implications for design in practice, both during education at IDE and afterwards in industry. First, an additional nuance to the conceptual framework of the design process from section five is provided with regards to the function of the design propositions. Second, conceptual descriptions of the design process are translated to the practical implementation of the design process in organizations, highlighting both the different types of organizational value of design, as well as the organizational functions from which design activities can be performed. Finally, the future of the Strategic Product Design (SPD) master is discussed, providing a perspective on how it could be developed towards the future.

---

### 12.1 Design propositions: to learn for yourself or for others?

In design we are used to pay attention to the thing that we are creating; the physical or imaginative artifact in front of us. However, based on the theoretical contemplations from the previous sections, I would like to argue that design is about a process of self enhancement, where externalizations of an individual's thoughts, that we are used to look at as 'designs' and 'products', are actually materials that are created to allow the individual to better understand the situation he or she is dealing with. We create learning materials that cannot be consumed anywhere else, as these artifacts are unique in the dimensions they represent to us. To design is to create new steppingstones in the process of trying to solve a problem that nobody has yet been able to find the right words and dimensions for to describe. Our interpretation of these objects, in the form of instantiated schemata, are the gateway for new information to intersect with existing instantiated representations of the world and their underlying logical structures.

I argue that design can be regarded as a process of *propositional learning*, where the designs and products are the offshoots of the learning process that allow both the designer to learn and to enhance his knowledge, as well as (an) external stakeholder(s). The act of creating design outputs that allow the external stakeholder(s) to learn can be regarded as value-creation. For example, a certain output as part of the problem analysis iteration might represent *personal enhancement*-oriented creation, whereas an output

generated as part of idea generation can be more focused on contributing to *value-creation* (figure 50). Both perspectives are important and should be aligned. From the perspective of the *value-creation*, it can be said that when one aims to design products for maximum user-value, it is relevant to look for ways in which one can improve one's ability to learn during the process of design. From the *learning* perspective, when one aims to extract relevant insights during the creation of products, one has to produce propositions that, at the time of creation, are expected to represent maximum value to a certain set of stakeholders. Here, 'relevant' is relative and depends on one's personal interests or other external incentives. Afterwards, the interaction between the proposition and the designer or stakeholders may present new dimensions or wrong assumptions that enhance the *learning* process and as such benefits both the *personal enhancement* of the designer and *value-creation* for the external stakeholder(s). Here the generated products can be anything ranging from spoken or written language, to drawings, images, renders or other representations of knowledge and the user represents any type of stakeholder.

The type of design that is taught within the bachelor and master programs of the Industrial Design Engineering (IDE) discipline at the TU Delft can be described as a process of propositional learning where one's goals surrounding personal learning or enhancement are aligned with the needs, desires and other incentives presented by external stakeholders such that the learning process is directed at solving a problem of a specific stakeholder. An industrial designer makes another's problem his own problem.

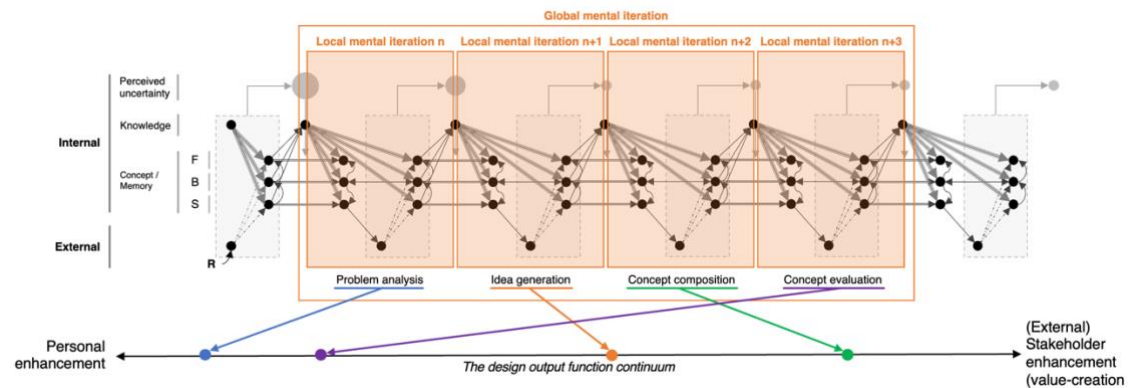


Figure 50: The continuum of design output functions

## 12.2 From design in theory to design in practice

This section aims to detail how different forms of design can be embedded in different types of organizations. This is relevant for students who have finished their masters in one of the three masters of the Industrial Design Engineering (IDE) Faculty of the TU Delft, and who are confronted with the choice between working in the academic world or working in industry and its many different options. Moreover, it is important for the positioning of the design practice relative to other disciplines. The remainder of this section focuses on the design practice in industry as this is the most common place for

designers to go to after their studies<sup>31</sup>. First, a reflection is provided on how the different forms of design outputs contribute to organizations. Second, the relation between the three IDE masters, and specifically the Strategic Product Design (SPD) master, and organizational function is detailed.

### ***12.2.1 The organizational value of design***

Concluding from the previous section, the industrial design process is inherently a combination of *learning* and ultimately *value-creation*. The implications of this observation are critical for the materialization and integration of the design process in practice. Attending to the two types of design output as defined in the previous section, it can be argued that design provides two types of contributions to a value-creation oriented organization.

#### ***1. Contributing by progressing the understanding of the status quo***

Here, the design process uses a learning-first approach to provide new information or perspectives that represent the fundamental information infrastructure for determining, among others, how an organization can (re)structure, (re)position, (re)brand, what type of products or services it should develop, and how these new products can be developed given the organizational/technological/societal landscape. Often, this type of design contribution happens in the context of uniting issues from different domains under one big, and often ambiguous, research area, with the objective to build towards a more common ground (and thus less siloed) approach to solving issues.

#### ***2. Contributing by activating (existing) knowledge***

In this instance, the design process focusses on translating existing knowledge into propositions that represent value to a certain group of stakeholders. In the case of the board or management of an organization, the value proposition can be a strategic roadmap. When the stakeholders are customers, the proposition can take the form of certain products and services. Both types of design-based contributions often happen in tandem, positioned in iterative learning cycles.

During *learning*, the design process focusses on generating artifacts that maximize the potential for learning and hence the potential for breakthrough insights. During *value-creation*, the design process focusses on converging towards actionable tasks and deliverables that represent value to surrounding stakeholders. These tasks and deliverables still represent the learning materials at a specific point in time, but at that point in time their objective lies closer to delivering value than initiating another round

---

<sup>31</sup> Based on the observation of the lower number of people pursuing an academic career in the form of a PhD after their IDE bachelor's and master's degrees, relative to those who are going to work for other public and private organizations.

of reflection and learning. Again, these design outputs have to be interpreted as being positioned somewhere on the continuum of design output functions (figure 50).

### ***12.2.2 The design process related to organizational functions***

An important observation is that, although there is no such thing as black and white, the function of the output of the two types of design contributions are different in nature and hence have a different relation with the organization. The first type focusses on building towards a better infrastructure for future organizational pursuits. It is strategic in nature and, depending on the organization, is either outsourced (publicly: academia, privately: consultancies) or organized in-house within an R&D department or at management level. The second type is about generating output that, depending on the type of output, has the potential to deliver more short-term business returns. As such, it can be described as an activity that is situated closer to the core business of an organization. However, an example of a design output that is still not directly linked to the core business of an organization (except for academia and consultancies in some scenarios), is a strategic roadmap.

### ***12.2.3 Strategic Product Design master related to organizational functions***

The Strategic Product Design (SPD) master covers both types of design-based organizational contributions (section 11.2.1), focusing mostly on the strategic aspects related to the future business of the organization. Although SPD educates designers who are capable of both doing the problem analysis as well as the creation of propositions, these propositions often take the form of strategic documents or visions that are not directly contributing to short term organizational value-creation. This leads to an important remark about the position from which strategic design is executed: Strategic design is a supportive function relative to the purpose of the organization, since there is no direct connection between this type of design activity and the core business of an organization (except when the organization's core business is provide advice on organizational strategy). In other words, strategic design is either an external advisory or internal corporate function. As such, the core expertise of SPD in its current form is best applied in academic, management, or consultancy functions.

## **12.3 The future of Strategic (product) Design**

The previous section has shown that strategic design happens at either the (or a specific) management level within an organization or from a position in an external advisory or consulting organizations. Although the link between the design process and organizational functions and the specific strategic of them instances are clarified, it is considered hard by SPD students to express the value of their educational background for the organization<sup>32</sup>. To build towards a better definition of SPD, (self-reported)

---

<sup>32</sup> This is based on my own experiences, and those of other SPD students that I have talked to over past two years.

challenges of SPD students regarding their value and professional identity are discussed in the light of the observations from the previous sections. Afterwards, the implications for the future of the SPD master are discussed.

### ***12.3.1 Addressing the ‘challenges of an SPD’er’***

Reflecting on my own experience and those of others in explaining what SPD is and what value it represents, there have been three main challenges. Below, these challenges are listed, including a reflection on how they can be (partially) addressed based on the remarks presented above.

#### ***1. It has been difficult to position SPD relative to other disciplines.***

A way to position SPD is to argue that strategic design is a combination of holistic problem analysis (*learning*) and proposition creation (*value-creation*) executed on a strategic level of abstraction in the context of either research or management-level decisions. Here, ‘holistic’ stresses the integrated nature of the design approach, therefore distancing design from problem solving in engineering disciplines that take an approach that favors the break-down of large problems into smaller, more easily solvable problems. The word ‘proposition’ is a more neutral expression of ‘product’, while simultaneously comprising more direction than another neutral word ‘artifact’. This definition may need a bit of adjustment for it to be applicable in the public domain due to different organizational structures.

#### ***2. The words ‘product’ and ‘design’ have a vary ambiguous meaning, often leading to confusion rather than clarity.***

To address this issue, the mix of words in the name ‘strategic product design’ has to be reconsidered. Based on my own experience, the word ‘design’ can remain the same, only under the condition that the word ‘product’ is eliminated. For example, when combining ‘product’ and ‘design’, people’s associations often lead to the interpretation of SPD as being a study on physical products or UI/UX design. When the word ‘product’ is eliminated, design becomes a more abstract term that can be more freely associated with the creation of roadmaps (as in the course ‘Design Roadmapping’), policy, or other strategic assets.

#### ***3. (Building on the first challenge) It has been hard to make the organization-specific value-case of strategic design.***

This third challenge is the most difficult to address, as it requires the most abstract translation from design process-related skills to the specific domain of an organization or institution. Moreover, it presents a dimension that is related to the two types of design output functions and requires more attention in IDE’s education: domain-specific knowledge (immediate stakeholder value creation) versus general process-specific expertise (personal enhancement in line with certain (internally or externally defined) learning objectives). Organizations that provide advice or consultation might be more open to adopt individuals that have expertise in the domain of the design process

without further industry- or domain-specific knowledge. The nature of the work at these organizations is often defined by the help their clients need with regards to specific changes in industry or market conditions. This type of activity requires both parties to learn, and hence to build the knowledge infrastructure for the future course of action. Here, domain-specific knowledge may be helpful, but also partially subordinate to one's ability to learn and adapt. However, non-consultancy organizations might be looking for more domain-specific experts that don't have to learn anymore. If 'learning' is not central to the job or organization, and the focus is instead on executing (hence value-creation) one's existing knowledge becomes very important as the conditions in which one is expected to start operating may require a very constant and predictable input, something that a learning process cannot deliver. The only way an SPD student can be positioned strongly towards the latter type of organizations is by 1) getting lucky with the type of client cases that are offered during the master curriculum, as these cases may be positioned within a domain of interest of the organization, or 2) by pursuing an internship that is positioned in a relevant domain. This third challenge requires either a stronger top-down marketing narrative of the IDE discipline, or bottom-up changes in the curriculum.

### ***12.3.2 Implications for the future of the SPD master***

Building on the above-listed reflection on SPD-related challenges, it can be stated that future improvements of the SPD master can be made in the domains of 1) finding the right words to describe the relative position of the design practice, and 2) establishing a good balance between design process-specific expertise and domain-specific knowledge. Two routes can be established.

#### ***Route 1: Focus on process-related expertise***

In this scenario, SPD is regarded as a mostly process-related skillset that is more closely positioned to management-related studies. Moreover, SPD can be renamed into Strategic Design & Business/Public Administration (SD&B/PM), as the word 'product' does not fit the context of process-oriented propositions anymore and the understanding of organizational process management becomes more central. A distinction can then be made between public and private sector tracks given the different types of organizational dynamics. This SD&B/PM master may be integrated with courses from Technology, Policy & Management, to not only increase the consolidation of the design practice in the public domain, but to also better integrate policy design with the conventions around the design of products and services.

#### ***Route 2: A hybrid between process expertise and domain-knowledge***

If strategic design is considered a combination of both domain-specific and process-related expertise, its position as a discipline would be weaker, although the potential for reaching non-consultancy-like job positions would be higher. Another implication is that there should be more domain-specific educational tracks to better align job-



expectations with academic knowledge. In this route, the name of the master remains an open question as there is neither a specific focus on the product or the process.

Combining both routes with the distinction between design outputs as learning or value-creation materials (figure 50), the quadrant as shown in figure 51 can be constructed. This quadrant shows three aspects. First, it highlights the content of the two routes divided over the ‘personal learning’ and ‘value-creation’ aspects of the design process. Second, it shows the type of job position that a strategic designer can end up in, visualized by the four dots that represent the four combinations of both axes. Finally, the type of institutions and organizations that a strategic designer can work for, indicated by the color-coded squares in each quadrant, are provided.

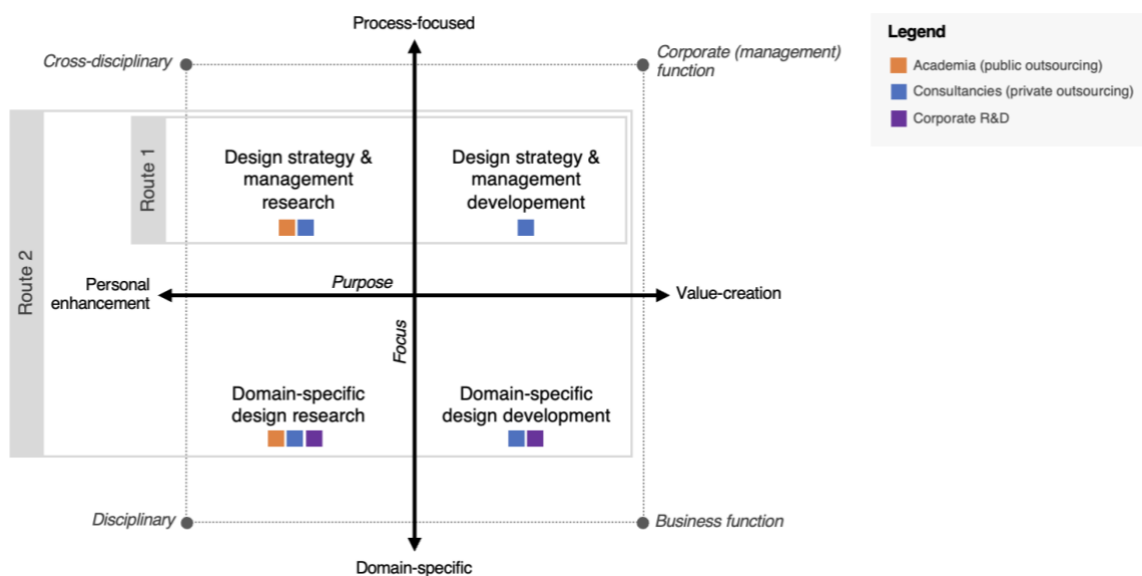


Figure 51: A visual overview of the aspects related to the future direction of the SPD master.

### 12.3.3 My personal perspective

Based on my experience, I think the future of SPD should align with route 1. Building on this direction, I think Strategic Design & Business/Public Administration (SD&B/PM) should be positioned in between Business Administration, Management and engineering (figure 52). Within this niche, SD&B/PM is able to bridge various domains of knowledge through a process-enabled learning approach, where this learning process is directed and informed by the learning goals of both the designer and external stakeholders to ensure value-creation along the learning curve.

Compared to Business Administration and Management, SD&B/PM should have an edge when it comes to expertise about aligning personal and external learning objectives to build things that represent value. Moreover, SD&B/PM should have enough technical expertise to unite domain-specific expertise with organizational

processes. The result is that SD&B/PM can be positioned as a discipline that enables individuals to more proactively shape the creation of holistically integrated propositions. This is very valuable in the light of the big future challenges that already present very ambiguous, complex, and multidimensional scenarios in which a lot of learning is required. In short, SD&B/PM is management with imagination about how to build domain-specific-knowledge-exceeding concepts that create value.

Compared to engineering studies, SD&B/PM can be positioned as an alternative to systems-engineering with an additional focus on organizational objectives. Moreover, SD&B/PM should have an edge in scenarios when ambiguous problems have to be united in a holistic multi-disciplinary approach.

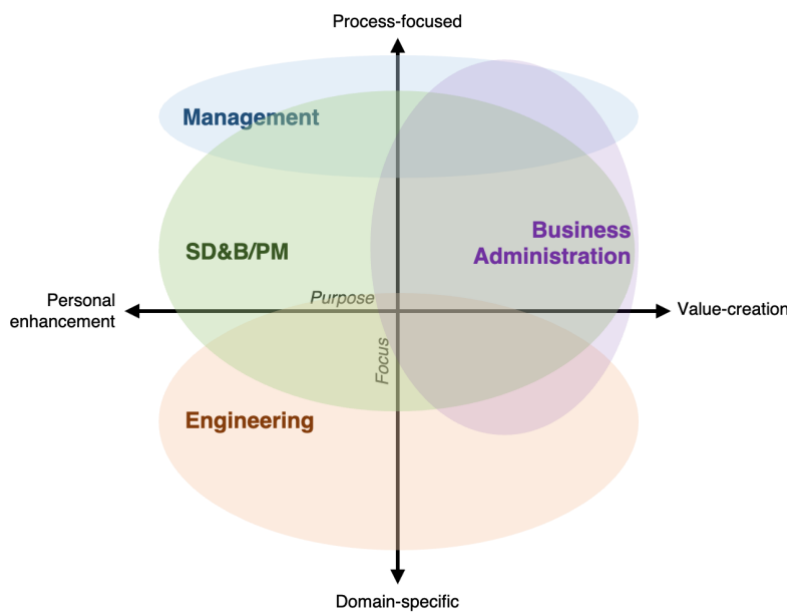


Figure 52: Relative positioning of strategic design

Strategic Design & Business/Public Administration, as positioned in figure 52, has the potential to replace a purely management-oriented approach within organizations. Instead of following a MECE (Mutually Exclusive, Collectively Exhaustive) regime by separating management expertise and domain-specific knowledge, SD&B/PM can introduce a more nuanced and fluid transition between management and operational functions to potentially establish a more resilient form of organizational operations. In the light of increasing geo-political tensions and the environment-driven changes in the future policy-landscape, such resilience might be the key ingredient for organizations to adapt swiftly and effectively.

## 13. Bibliography

- Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association: JMLA*, 103(3), 152–153. <https://doi.org/10.3163/1536-5050.103.3.010>
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, 45(2), 357–376. <https://doi.org/10.1037/0022-3514.45.2.357>
- Banjade, R., Oli, P., Sajib, M. I., & Rus, V. (2024). Identifying Gaps in Students' Explanations of Code Using LLMs. In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial Intelligence in Education* (Vol. 14830, pp. 268–275). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-64299-9\\_21](https://doi.org/10.1007/978-3-031-64299-9_21)
- Barron, F. (1995). *No rootless flower: An ecology of creativity*. Cresskill, NJ: Hampton Press.
- Berni, A., Borgianni, Y., Rotini, F., Gonçalves, M., & Thoring, K. (2024). Stimulating design ideation with artificial intelligence: Present and (short-term) future. *Proceedings of the Design Society*, 4, 1939–1948. <https://doi.org/10.1017/pds.2024.196>
- Bessemer, S. P., & Treffinger, D. J. (1981). Analysis of Creative Products: Review and Synthesis\*. *The Journal of Creative Behavior*, 15(3), 158–178. <https://doi.org/10.1002/j.2162-6057.1981.tb00287.x>
- Bessemer, S. P. and K. O'Quin (1987) *Creative Product Analysis: Testing a Model by Developing a Judging Instrument*" in S. G. Isaksen (ed.) *Frontiers of Creativity: Beyond the Basics*, Buffalo, New York: Bearly.
- Boden, M. A. (1998). Creativity and artificial intelligence.
- Boden, M. A. (2009). Computer Models of Creativity. *AI Magazine*, 30(3), 23–34. <https://doi.org/10.1609/aimag.v30i3.2254>
- Bostrom, N. (2012). The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents. *Minds and Machines*, 22(2), 71–85. <https://doi.org/10.1007/s11023-012-9281-3>
- Bown, O. (2015). Player Responses to a Live Algorithm: Conceptualising computational creativity without recourse to human comparisons?
- Campbell, D. T. (1962). Blind variation and selective retention in creative thought as in other knowledge processes. *General System Yearbook*, 17.
- Cash, P., Gonçalves, M., & Dorst, K. (2023). Method in their madness: Explaining how designers think and act through the cognitive co-evolution model. *Design Studies*, 88, 101219. <https://doi.org/10.1016/j.destud.2023.101219>
- Cash, P., & Kreye, M. (2017). Uncertainty Driven Action (UDA) model: A foundation for unifying perspectives on design activity. *Design Science*, 3, e26. <https://doi.org/10.1017/dsj.2017.28>
- Chen, L., Tsang, Y., Jing, Q., & Sun, L. (2024). A LLM-augmented Morphological Analysis Approach for Conceptual Design. *DRS2024: Boston*. <https://doi.org/10.21606/drs.2024.605>
- Christensen, B. T., & Ball, L. J. (2018). Fluctuating epistemic uncertainty in a design team as a metacognitive driver for creative cognitive processes. *CoDesign*, 14(2), 133–152. <https://doi.org/10.1080/15710882.2017.1402060>

- Couger, J. D., Higgins, L. F., & McIntyre, S. C. (1993). (Un)Structured Creativity in Information Systems Organizations. *MIS Quarterly*, 17(4), 375.  
<https://doi.org/10.2307/249584>
- Court, A. W. (1997). The relationship between information and personal knowledge in new product development. *International Journal of Information Management*, 17(2), 123–138. [https://doi.org/10.1016/S0268-4012\(96\)00054-0](https://doi.org/10.1016/S0268-4012(96)00054-0)
- Crilly, N. (2021a). The Evolution of “Co-evolution” (Part I): Problem Solving, Problem Finding, and Their Interaction in Design and Other Creative Practices. *She Ji: The Journal of Design, Economics, and Innovation*, 7(3), 309–332.  
<https://doi.org/10.1016/j.sheji.2021.07.003>
- Crilly, N. (2021b). The Evolution of “Co-evolution” (Part II): The Biological Analogy, Different Kinds of Co-evolution, and Proposals for Conceptual Expansion. *She Ji: The Journal of Design, Economics, and Innovation*, 7(3), 333–355.  
<https://doi.org/10.1016/j.sheji.2021.07.004>
- Crilly, N., & Moroşanu Firth, R. (2019). Creativity and fixation in the real world: Three case studies of invention, design and innovation. *Design Studies*, 64, 169–212.  
<https://doi.org/10.1016/j.destud.2019.07.003>
- Cropley, A. (2006). Creativity: A social approach. *Roeper Review*, 28(3), 125–130.  
<https://doi.org/10.1080/02783190609554351>
- Davis, N., Hsiao, C.-P., Popova, Y., & Magerko, B. (2015). An Enactive Model of Creativity for Computational Collaboration and Co-creation. In N. Zagalo & P. Branco (Eds.), *Creativity in the Digital Age* (pp. 109–133). Springer London.  
[https://doi.org/10.1007/978-1-4471-6681-8\\_7](https://doi.org/10.1007/978-1-4471-6681-8_7)
- Dawkins, R. (1976). *The selfish gene*. Oxford university press, 189–201.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507. <https://doi.org/10.1007/s11097-007-9076-9>
- Dean, D., Hender, J., Henley Management College, Rodgers, T., Consultant College Station, Texas, Santanen, E., & Bucknell University. (2006). Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation. *Journal of the Association for Information Systems*, 7(10), 646–699.  
<https://doi.org/10.17705/1jais.00106>
- Deci, E. L., & Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. Springer US. <https://doi.org/10.1007/978-1-4899-2271-7>
- Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., & Ebel, P. (2019). The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems. *Hawaii International Conference on System Sciences*.  
<https://doi.org/10.24251/HICSS.2019.034>
- Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business & Information Systems Engineering*, 61(5), 637–643.  
<https://doi.org/10.1007/s12599-019-00595-2>
- Dewey, J. (1934). *Art as experience* (Penguin).
- Döppner, D. A., Derckx, P., & Schoder, D. (2019). Symbiotic Co-Evolution in Collaborative Human-Machine Decision Making: Exploration of a Multi-Year Design Science Research Project in the Air Cargo Industry. *Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2019.033>
- Dorst, K. (2019). Co-evolution and emergence in design. 65.

- Dorst, K., & Cross, N. (2001). Creativity in the design process: Co-evolution of problem–solution. *Design Studies*, 22(5), 425–437. [https://doi.org/10.1016/S0142-694X\(01\)00009-6](https://doi.org/10.1016/S0142-694X(01)00009-6)
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Engelbart, D. C. (1962). Augmenting human intellect: A conceptual framework. In *Augmented Education in the Global Age* (pp. 13-29). Routledge.
- Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–459. <https://doi.org/10.1016/j.tics.2003.08.012>
- Fabri, L., Häckel, B., Oberländer, A. M., Rieg, M., & Stohr, A. (2023). Disentangling Human-AI Hybrids: Conceptualizing the Interworking of Humans and AI-Enabled Systems. *Business & Information Systems Engineering*, 65(6), 623–641. <https://doi.org/10.1007/s12599-023-00810-1>
- Franceschelli, G., & Musolesi, M. (2023). On the Creativity of Large Language Models (arXiv:2304.00008). arXiv. <http://arxiv.org/abs/2304.00008>
- Fyhn, B., Schei, V., & Sverdrup, T. E. (2023). Taking the emergent in team emergent states seriously: A review and preview. *Human Resource Management Review*, 33(1), 100928. <https://doi.org/10.1016/j.hrmr.2022.100928>
- Gero, J. S. (1990). Design Prototypes: A Knowledge Representation Schema for Design.
- Gero, J. S., & Kannengiesser, U. (2004). The situated function–behaviour–structure framework. *Design Studies*, 25(4), 373–391. <https://doi.org/10.1016/j.destud.2003.10.010>
- Gero, J. S., Kannengiesser, U., & Crilly, N. (2022). Abstracting and formalising the design co-evolution model. *Design Science*, 8, e14. <https://doi.org/10.1017/dsj.2022.10>
- Glăveanu, V. P. (2013). Rewriting the Language of Creativity: The Five A’s Framework. *Review of General Psychology*, 17(1), 69–81. <https://doi.org/10.1037/a0029528>
- Glăveanu, V. P. (2015). From Individual Agency to Co-agency. In C. W. Gruber, M. G. Clark, S. H. Klempe, & J. Valsiner (Eds.), *Constraints of Agency* (pp. 245–265). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10130-9\\_15](https://doi.org/10.1007/978-3-319-10130-9_15)
- Goel, A. K., & Bhatta, S. R. (2004). Use of design patterns in analogy-based design. *Advanced Engineering Informatics*, 18(2), 85–94. <https://doi.org/10.1016/j.aei.2004.09.003>
- Grimm, S. R. (2014). Understanding as knowledge of causes. In *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 329-345). Cham: Springer International Publishing.
- Grossen, M. (2008). Methods for studying collaborative creativity: An original and adventurous blend. *Thinking Skills and Creativity*, 3(3), 246–249. <https://doi.org/10.1016/j.tsc.2008.09.005>
- Grossman, R., Friedman, S. B., & Kalra, S. (2017). Teamwork Processes and Emergent States. In E. Salas, R. Rico, & J. Passmore (Eds.), *The Wiley Blackwell Handbook of the*

- Psychology of Team Working and Collaborative Processes (1st ed., pp. 243–269). Wiley. <https://doi.org/10.1002/9781118909997.ch11>
- Halstrøm, P. L., & Galle, P. (2015). Design as Co-Evolution of Problem, Solution, and Audience. *Artifact*, 3(4), 3. <https://doi.org/10.14434/artifact.v3i4.12815>
  - Hao, S., Gu, Y., Ma, H., Hong, J. J., Wang, Z., Wang, D. Z., & Hu, Z. (2023). Reasoning with Language Model is Planning with World Model (arXiv:2305.14992). arXiv. <http://arxiv.org/abs/2305.14992>
  - Hatchuel, A., & Weil, B. (2002). Notions and applications of a unified design theory.
  - Heer, J. (2019). Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116(6), 1844–1850. <https://doi.org/10.1073/pnas.1807184115>
  - Horn, D., & Salvendy, G. (2006). Consumer-based assessment of product creativity: A review and reappraisal. *Human Factors and Ergonomics in Manufacturing*, 16(2), 155–175. <https://doi.org/10.1002/hfm.20047>
  - Jackson, P. W. and S. Messick (1965) "The Person, the Product, and the Response: Conceptual Problems in the Assessment of Creativity," *Journal of Personality and Social Psychology* (33) 3, pp. 309-329.
  - Jiang, X., Tian, Y., Hua, F., Xu, C., Wang, Y., & Guo, J. (2024). A Survey on Large Language Model Hallucination via a Creativity Perspective (arXiv:2402.06647). arXiv. <http://arxiv.org/abs/2402.06647>
  - Jordanous, A. (2016). Four PPPerspectives on computational creativity in theory and in practice. *Connection Science*, 28(2), 194–216. <https://doi.org/10.1080/09540091.2016.1151860>
  - Kagan J. ed. (1967). *Creativity and learning*. Boston, MA: Beacon Press Press.
  - Kalyan, K. S. (2024). A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Natural Language Processing Journal*, 6, 100048. <https://doi.org/10.1016/j.nlp.2023.100048>
  - Kantosalo, A., & Takala, T. (2020). Five C's for Human–Computer Co-Creativity—An Update on Classical Creativity Perspectives.
  - Kaufman, J. C., & Sternberg, R. J. (Eds.). (2010). *The Cambridge handbook of creativity*. Cambridge University Press.
  - Kim, J., & Maher, M. L. (2023). The effect of AI-based inspiration on human design ideation. *International Journal of Design Creativity and Innovation*, 11(2), 81–98. <https://doi.org/10.1080/21650349.2023.2167124>
  - Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521–3526. <https://doi.org/10.1073/pnas.1611835114>
  - Langrish, J. Z. (2004). Darwinian Design: The Memetic Evolution of Design Ideas.
  - Lawton, T., Grace, K., & Ibarrola, F. J. (2023). When is a Tool a Tool? User Perceptions of System Agency in Human–AI Co-Creative Drawing. *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, 1978–1996. <https://doi.org/10.1145/3563657.3595977>
  - Lawton, T., Ibarrola, F. J., Ventura, D., & Grace, K. (2023). Drawing with Reframer: Emergence and Control in Co-Creative AI. *Proceedings of the 28th International*



- Conference on Intelligent User Interfaces, 264–277.  
<https://doi.org/10.1145/3581641.3584095>
- Leonardi, P. M. (2013). Theoretical foundations for the study of sociomateriality. *Information and Organization*, 23(2), 59–76.  
<https://doi.org/10.1016/j.infoandorg.2013.02.002>
  - Li, J., & Li, J. (2024). A Map of Exploring Human Interaction patterns with LLM: Insights into Collaboration and Creativity (arXiv:2404.04570). arXiv.  
<http://arxiv.org/abs/2404.04570>
  - Lin, Z., Ehsan, U., Agarwal, R., Dani, S., Vashishth, V., & Riedl, M. (2023). Beyond Prompts: Exploring the Design Space of Mixed-Initiative Co-Creativity Systems (arXiv:2305.07465). arXiv. <http://arxiv.org/abs/2305.07465>
  - Liu, Y., Chen, S., Cheng, H., Yu, M., Ran, X., Mo, A., Tang, Y., & Huang, Y. (2024). How AI Processing Delays Foster Creativity: Exploring Research Question Co-Creation with an LLM-based Agent. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–25. <https://doi.org/10.1145/3613904.3642698>
  - Liu, Z., Zhang, Y., Li, P., Liu, Y., & Yang, D. (2023). Dynamic LLM-Agent Network: An LLM-agent Collaboration Framework with Agent Team Optimization (arXiv:2310.02170). arXiv. <http://arxiv.org/abs/2310.02170>
  - Lyytinen, K., & Newman, M. (2008). Explaining information systems change: A punctuated socio-technical change model. *European Journal of Information Systems*, 17(6), 589–613. <https://doi.org/10.1057/ejis.2008.50>
  - Maher, M. L. (1994). Creative design using a genetic algorithm. In *Computing in Civil Engineering* (pp. 2014–2021). ASCE.
  - Maher, M. L., & Poon, J. (1995, November). Co-evolution of the fitness function and design solution for design exploration. In *Proceedings of 1995 IEEE International Conference on Evolutionary Computation* (Vol. 1, p. 240). IEEE.
  - Maher, M. L., & Poon, J. (1996). Modeling Design Exploration as Co-Evolution.
  - Mao, S., Zhang, N., Wang, X., Wang, M., Yao, Y., Jiang, Y., Xie, P., Huang, F., & Chen, H. (2023). Editing Personality for LLMs (arXiv:2310.02168). arXiv.  
<http://arxiv.org/abs/2310.02168>
  - Marks, M. A., John E., M., & Stephen J., Z. (2001). A Temporally Based Framework and Taxonomy of Team Processes.
  - Markus, H. R., & Hamedani, M. G. (2007). Sociocultural psychology. *Handbook of cultural psychology*, 3–39.
  - Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality (Version 1). arXiv.  
<https://doi.org/10.48550/ARXIV.1310.4546>
  - Musker, S., Duchnowski, A., Millière, R., & Pavlick, E. (2024). Semantic Structure-Mapping in LLM and Human Analogical Reasoning (arXiv:2406.13803). arXiv.  
<http://arxiv.org/abs/2406.13803>
  - Nagasundaram, M. & Bostrom, R. P. (1994-95). "The Structuring of Creative Processes Using GSS: A Framework for Research," *Journal of Management Information Systems* (11) 3, pp. 87-114.
  - Norman, D. A., & Rumelhart, D. E. (1975). Explorations in Cognition. *The American Journal of Psychology*, 88(4), 691. <https://doi.org/10.2307/1421908>
  - OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom,

- V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., ... Zoph, B. (2024). GPT-4 Technical Report (arXiv:2303.08774). arXiv. <http://arxiv.org/abs/2303.08774>
- Organisciak, P., Acar, S., Dumas, D., & Berthiaume, K. (2023). Beyond semantic distance: Automated scoring of divergent thinking greatly improves with large language models. *Thinking Skills and Creativity*, 49, 101356. <https://doi.org/10.1016/j.tsc.2023.101356>
  - Orlikowski, W. J., & Scott, S. V. (2008). 10 Sociomateriality: Challenging the Separation of Technology, Work and Organization. *Academy of Management Annals*, 2(1), 433–474. <https://doi.org/10.5465/19416520802211644>
  - Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why Isn't Creativity More Important to Educational Psychologists? Potentials, Pitfalls, and Future Directions in Creativity Research. *Educational Psychologist*, 39(2), 83–96. [https://doi.org/10.1207/s15326985ep3902\\_1](https://doi.org/10.1207/s15326985ep3902_1)
  - Qin, C., Xia, W., Wang, T., Jiao, F., Hu, Y., Ding, B., Chen, R., & Joty, S. (2024). Relevant or Random: Can LLMs Truly Perform Analogical Reasoning? (arXiv:2404.12728). arXiv. <http://arxiv.org/abs/2404.12728>
  - Rezwana, J., & Maher, M. L. (2023). Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *ACM Transactions on Computer-Human Interaction*, 30(5), 1–28. <https://doi.org/10.1145/3519026>
  - Rhodes, M. (1961). An Analysis of Creativity. *The Phi Delta Kappan*, 42(7), 305–310. JSTOR.
  - Rick, S. R., Giacomelli, G., Wen, H., Laubacher, R. J., Taubenslag, N., Heyman, J. L., Knicker, M. S., Jeddi, Y., Maier, H., Dwyer, S., Ragupathy, P., & Malone, T. W. (2023). Supermind Ideator: Exploring generative AI to support creative problem-solving (arXiv:2311.01937). arXiv. <http://arxiv.org/abs/2311.01937>
  - Ritchie, G. (2001). Assessing creativity. *Proc. of AISB'01 Symposium*.
  - Ritchie, G. (2007). Some Empirical Criteria for Attributing Creativity to a Computer Program. *Minds and Machines*, 17(1), 67–99. <https://doi.org/10.1007/s11023-007-9066-2>
  - Rumelhart, D. E., & Ortony, A. (2017). The Representation of Knowledge in Memory 1. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the Acquisition of Knowledge* (1st ed., pp. 99–135). Routledge. <https://doi.org/10.4324/9781315271644-10>
  - Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, 55(1), 657–687. <https://doi.org/10.1146/annurev.psych.55.090902.141502>
  - Runco, M. A. (2023a). AI can only produce artificial creativity. *Journal of Creativity*, 33(3), 100063. <https://doi.org/10.1016/j.yjoc.2023.100063>
  - Runco, M. A. (2023b). Updating the Standard Definition of Creativity to Account for the Artificial Creativity of AI. *Creativity Research Journal*, 1–5. <https://doi.org/10.1080/10400419.2023.2257977>
  - Runco, M. A., & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, 24(1), 92–96. <https://doi.org/10.1080/10400419.2012.650092>
  - Sääksjärvi, M., & Gonçalves, M. (2018). Creativity and meaning: Including meaning as a component of creative solutions. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 32(4), 365–379. <https://doi.org/10.1017/S0890060418000112>



- Sawyer, R. K. (1998). Introduction. In R. K. Sawyer (Ed.), *Creativity in performance* (pp. 1–6). Greenwich, CT: Ablex.
- Sawyer, R. K., John-Steiner, V., Moran, S., Sternberg, R. J., Feldman, D. H., Nakamura, J., & Csikszentmihalyi, M. (2003). *Creativity and Development*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195149005.001.0001>
- Schon, D. A., & Wiggins, G. (1992). Kinds of seeing and their functions in designing. *Design Studies*, 13(2), 135–156. [https://doi.org/10.1016/0142-694X\(92\)90268-F](https://doi.org/10.1016/0142-694X(92)90268-F)
- Schwartz, S., Yaeli, A., & Shlomov, S. (2023). Enhancing Trust in LLM-Based AI Automation Agents: New Considerations and Future Challenges (arXiv:2308.05391). arXiv. <http://arxiv.org/abs/2308.05391>
- Shapiro, D., Li, W., Delaflor, M., & Toxtli, C. (2023). Conceptual Framework for Autonomous Cognitive Entities. <https://doi.org/10.13140/RG.2.2.14161.30569>
- Sharma, A., Rao, S., Brockett, C., Malhotra, A., Jojic, N., & Dolan, B. (2024). Investigating Agency of LLMs in Human-AI Collaboration Tasks.
- Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). Continual Learning with Deep Generative Replay (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.1705.08690>
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human–Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Simonton, D.K., 1999a. *Origins of genius: Darwinian perspectives on creativity*. Oxford, New York.
- Simonton, D.K., 1999b. Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry*, 10, 309–328.
- Smithers, T., Conkie, A., Doheny, J., Logan, B., Millington, K., & Tang, M. X. (1990). Design as intelligent behaviour: An AI in design research programme. *Artificial Intelligence in Engineering*, 5(2), 78–109. [https://doi.org/10.1016/0954-1810\(90\)90004-N](https://doi.org/10.1016/0954-1810(90)90004-N)
- Stevenson, C., Smal, I., Baas, M., & Grasman, R. (2022). Putting GPT-3’s Creativity to the (Alternative Uses) Test.
- Sun, F.-K., Ho, C.-H., & Lee, H.-Y. (2019). LAMOL: LAnguage MOdeling for Lifelong Language Learning. <https://doi.org/10.48550/ARXIV.1909.03329>
- Takeda, H., Veerkamp, P., & Yoshikawa, H. (1990). Modeling Design Process. *AI Magazine*, 11(4), 37–37. <https://doi.org/10.1609/aimag.v11i4.855>
- Tardif T. Z., & Sternberg R. J. (1988). What do we know about creativity? In R. J. Sternberg (Ed.), *The nature of creativity*, Chapter 17 (pp. 429–440). Cambridge, UK: Cambridge University Press.
- Taura, T., & Nagai, Y. (Eds.). (2010). *Design creativity 2010*. Springer Science & Business Media.
- Thompson, G., & Lordan, M. (1999). A review of creativity principles applied to engineering design. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 213(1), 17–31. <https://doi.org/10.1243/0954408991529960>
- Thoring, K., Huettemann, S., & Mueller, R. M. (2023). THE AUGMENTED DESIGNER: A RESEARCH AGENDA FOR GENERATIVE AI-ENABLED DESIGN. *Proceedings of the Design Society*, 3, 3345–3354. <https://doi.org/10.1017/pds.2023.335>

- Tsiakas, K., & Murray-Rust, D. (2024). Unpacking Human-AI interactions: From interaction primitives to a design space (arXiv:2401.05115). arXiv. <http://arxiv.org/abs/2401.05115>
- Tulving, E. (1989). Memory: Performance, knowledge, and experience. *European Journal of Cognitive Psychology*, 1(1), 3–26. <https://doi.org/10.1080/09541448908403069>
- van Eijck, J., & Schwarzentruher, F. (n.d.). Epistemic Probability Logic Simplified.
- Van Zoelen, E. M., Van Den Bosch, K., & Neerincx, M. (2021). Becoming Team Members: Identifying Interaction Patterns of Mutual Adaptation for Human-Robot Co-Learning. *Frontiers in Robotics and AI*, 8, 692811. <https://doi.org/10.3389/frobt.2021.692811>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). Attention Is All You Need (arXiv:1706.03762). arXiv. <http://arxiv.org/abs/1706.03762>
- Viros-i-Martin, A., & Selva, D. (2021a). A Framework to Study Human-AI Collaborative Design Space Exploration. Volume 6: 33rd International Conference on Design Theory and Methodology (DTM), V006T06A052. <https://doi.org/10.1115/DETC2021-67619>
- Viros-i-Martin, A., & Selva, D. (2021b). A Framework to Study Human-AI Collaborative Design Space Exploration. Volume 6: 33rd International Conference on Design Theory and Methodology (DTM), V006T06A052. <https://doi.org/10.1115/DETC2021-67619>
- Vosniadou, S., & Ortony, A. (Eds.). (1989). *Similarity and Analogical Reasoning* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511529863>
- Wang, S., Wei, Z., Choi, Y., & Ren, X. (2024). Can LLMs Reason with Rules? Logic Scaffolding for Stress-Testing and Improving LLMs (arXiv:2402.11442). arXiv. <http://arxiv.org/abs/2402.11442>
- Weinberger, A. B., Iyer, H., & Green, A. E. (2016). Conscious Augmentation of Creative State Enhances “Real” Creativity in Open-Ended Analogical Reasoning. *PLOS ONE*, 11(3), e0150773. <https://doi.org/10.1371/journal.pone.0150773>
- Weisberg, R. (1993). *Creativity: Beyond the myth of the genius*. New York, NY: W. H. Freeman.
- Wiltschnig, S., Christensen, B. T., & Ball, L. J. (2013). Collaborative problem–solution co-evolution in creative design. *Design Studies*, 34(5), 515–542. <https://doi.org/10.1016/j.destud.2013.01.002>
- Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A. H., White, R. W., Burger, D., & Wang, C. (2023). AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation (arXiv:2308.08155). arXiv. <http://arxiv.org/abs/2308.08155>
- Xia, C., & Maes, P. (2013). The design of artifacts for augmenting intellect. *Proceedings of the 4th Augmented Human International Conference*, 154–161. <https://doi.org/10.1145/2459236.2459263>
- Zelikman, E., Harik, G., Shao, Y., Jayasiri, V., Haber, N., & Goodman, N. D. (2024). Quiet-STaR: Language Models Can Teach Themselves to Think Before Speaking (arXiv:2403.09629). arXiv. <http://arxiv.org/abs/2403.09629>
- Zhou, C., Zhang, X., & Yu, C. (2023). How does AI promote design iteration? The optimal time to integrate AI into the design process. *Journal of Engineering Design*, 1–28. <https://doi.org/10.1080/09544828.2023.2290915>

# Appendices

## Appendix A: Graduation project proposal



7174

# IDE Master Graduation Project



**Project team, procedural checks and Personal Project Brief**

In this document the agreements made between student and supervisory team about the student's IDE Master Graduation Project are set out. This document may also include involvement of an external client, however does not cover any legal matters student and client (might) agree upon. Next to that, this document facilitates the required procedural checks:

- Student defines the team, what the student is going to do/deliver and how that will come about
- Chair of the supervisory team signs, to formally approve the project's setup / Project brief
- SSC E&SA (Shared Service Centre, Education & Student Affairs) report on the student's registration and study progress
- IDE's Board of Examiners confirms the proposed supervisory team on their eligibility, and whether the student is allowed to start the Graduation Project

**STUDENT DATA & MASTER PROGRAMME**  
 Complete all fields and indicate which master(s) you are in

Family name	Kortenbach	IDE master(s)	<input type="checkbox"/> IPD	<input type="checkbox"/> Dfi	<input checked="" type="checkbox"/> SPD
Initials	T.	2 <sup>nd</sup> non-IDE master			
Given name	Tom	Individual programme (date of approval)			
Student number		Medisign	<input type="checkbox"/>		
		HPM	<input type="checkbox"/>		

**SUPERVISORY TEAM**  
 Fill in the required information of supervisory team members. If applicable, company mentor is added as 2<sup>nd</sup> mentor

Chair	Rebecca Price	dept./section	Design, organization & strategy	<div style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;"> <p>! Ensure a heterogeneous team. In case you wish to include team members from the same section, explain why.</p> </div> <div style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;"> <p>! Chair should request the IDE Board of Examiners for approval when a non-IDE mentor is proposed. Include CV and motivation letter.</p> </div> <div style="border: 1px solid #ccc; padding: 5px;"> <p>! 2<sup>nd</sup> mentor only applies when a client is involved.</p> </div>
mentor	Alessandro Bozzon	dept./section	Sustainable design engineering	
2 <sup>nd</sup> mentor				
client:				
city:		country:		
optional comments				

**APPROVAL OF CHAIR on PROJECT PROPOSAL / PROJECT BRIEF -> to be filled in by the Chair of the supervisory team**

Sign for approval (Chair)

Name Rebecca Price

Date 22-05-2024

Signature 

**CHECK ON STUDY PROGRESS**

To be filled in by SSC E&SA (Shared Service Centre, Education & Student Affairs), after approval of the project brief by the chair. The study progress will be checked for a 2<sup>nd</sup> time just before the green light meeting.

Master electives no. of EC accumulated in total \_\_\_\_\_ EC

Of which, taking conditional requirements into account, can be part of the exam programme \_\_\_\_\_ EC

<input checked="" type="checkbox"/>	<b>YES</b>	all 1 <sup>st</sup> year master courses passed
<input type="checkbox"/>	<b>NO</b>	missing 1 <sup>st</sup> year courses

Comments:

Sign for approval (SSC E&SA)

Name Rik Ledoux

Date 23-05-2024

Signature 

**APPROVAL OF BOARD OF EXAMINERS IDE on SUPERVISORY TEAM -> to be checked and filled in by IDE's Board of Examiners**

Does the composition of the Supervisory Team comply with regulations?

<b>YES</b>	<input checked="" type="checkbox"/>	Supervisory Team approved
<b>NO</b>	<input type="checkbox"/>	Supervisory Team not approved

Comments:

Based on study progress, students is ...

<input checked="" type="checkbox"/>	<b>ALLOWED</b> to start the graduation project
<input type="checkbox"/>	<b>NOT</b> allowed to start the graduation project

Comments:

Sign for approval (BoEx)

Name Monique von Morgen

Date 28/5/2024

Signature 



## Personal Project Brief – IDE Master Graduation Project

Name student **Tom Kortebach**

Student number  

### PROJECT TITLE, INTRODUCTION, PROBLEM DEFINITION and ASSIGNMENT

Complete all fields, keep information clear, specific and concise

**Project title** Exploring human-LLM interactions in co-design through the proposition of a new digital interface

*Please state the title of your graduation project (above). Keep the title compact and simple. Do not use abbreviations. The remainder of this document allows you to define and clarify your graduation project.*

### Introduction

*Describe the context of your project here; What is the domain in which your project takes place? Who are the main stakeholders and what interests are at stake? Describe the opportunities (and limitations) in this domain to better serve the stakeholder interests. (max 250 words)*

Context: Large Language Models (LLMs) have shown ever-increasing capabilities, such as multi-step reasoning and growing context windows. Research into LLMs is highly multi-disciplinary. LLMs have garnered attention in the field of design thinking (Asadi, 2023) and have prompted researchers with the question of how to integrate the LLM in the design ideation and validation phases (Tholander & Jonsson, 2023). Their empirical research has conceptual overlap with that in the field of more-than-human design (Giaccardi & Redström, 2020; Wakkary, 2020), co-performance between human and machine (Kuijter & Giaccardi, 2018) and machine agency (Pickering, Engen & Walland, 2017), all moving towards the notion of creativity as a construct that is shared between human and computational actors.

Another relevant field of research surrounding LLMs is that of creativity, as creativity and design are highly related and can both be studied through the lens of the other (Lazar, 2018)(Taura 7 Nagai, 2010). Subfields include computational creativity (from computer science) and design creativity (from design). Within computational creativity, the field of human-computer co-creativity focusses on the collaboration between human and machine in creativity, representing an adjacent line of research next to that of the human-AI research in the field of design. Finally, related to the latter is the field of human-computer interaction (HCI) that focusses on defining patterns of interaction between humans and machines.

Domain: This graduation project sits at the intersection of design, creativity and computer science. More specifically, it looks at research from the fields of design creativity & processes, HCI and computational creativity through the lens of LLMs. From this perspective, this graduation will contribute to the field of design research.

Opportunity: Given these capabilities and characteristics of LLMs as human amplifiers and the evolving field of design methodology, requiring ever more and abstract knowledge as it moves through the 5 orders of design knowledge (Mortati, 2022), the exploration of LLMs is highly relevant for design researchers and practitioners.

→ space available for images / figures on next page

introduction (continued): space for images

**sources**

- Lazar, L. (2018). The Cognitive Neuroscience of Design Creativity. *Journal of Experimental Neuroscience*, 12, 117906951880966. <https://doi.org/10.1177/1179069518809664>

- Taura, T., & Nagai, Y. (Eds.). (2010). *Design creativity 2010*. Springer Science & Business Media.

- Asadi, A. R. (2023, September). LLMs in Design Thinking: Autoethnographic Insights and Design Implications. In *Proceedings of the 2023 5th World Symposium on Software Engineering* (pp. 55-60).

- Tholander, J., & Jonsson, M. (2023, July). Design ideation with ai-sketching, thinking and talking with Generative Machine Learning Models. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (pp. 1930-1940).

- Lenneke Kuijter and Elisa Giaccardi. 2018. Co-performance: Conceptualizing the Role of Artificial Agency in the Design of Everyday Life. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173699>

- J. Brian Pickering, Vegard Engen, and Paul Walland. 2017. The Interplay Between Human and Machine Agency. In *Human-Computer Interaction. User Interface Design, Development and Multimodality (Lecture Notes in Computer Science)*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 47–59. [https://doi.org/10.1007/978-3-319-58071-5\\_4](https://doi.org/10.1007/978-3-319-58071-5_4)

- Elisa Giaccardi, Johan Redström. Technology and More-Than-Human Design. *Design issues* 2020; 36 (4): 33–44. doi: [https://doi.org/10.1162/desi\\_a\\_00612](https://doi.org/10.1162/desi_a_00612)

- Ron Wakkary. 2020. Nomadic Practices: A Posthuman Theory for Knowing Design. *International Journal of Design* 14 (Dec. 2020), 117–128.

- Marzia Mortati. New Design Knowledge and the Fifth Order of Design. *Design Issues* 2022; 38 (4): 21–34. doi: [https://doi.org/10.1162/desi\\_a\\_00695](https://doi.org/10.1162/desi_a_00695)

image / figure 1



image / figure 2





## Personal Project Brief – IDE Master Graduation Project

### Problem Definition

*What problem do you want to solve in the context described in the introduction, and within the available time frame of 100 working days? (= Master Graduation Project of 30 EC). What opportunities do you see to create added value for the described stakeholders? Substantiate your choice.  
(max 200 words)*

Theoretical investigations into the creativity of LLMs highlight the importance of the interaction between the LLM with an environment including humans for the creative potential of the human and LLM. This aligns with findings in the field of research on human-computer co-creativity (Kantosalo & Takala, 2020). However, research has pointed at the limitations of the current chat-based interfaces for the interaction between LLM and designer and presents this as an area for future research (Asadi, 2023; Tholander & Jonsson, 2023). In short, although there is an emerging understanding of the relation between LLMs and humans in the domain of creativity and design, there is no dedicated interface that facilitates the application of LLMs in those domains.

- Kantosalo, A., & Takala, T. (2020). Five C's for Human-Computer Co-Creativity—An Update on Classical Creativity Perspectives.
- Asadi, A. R. (2023, September). LLMs in Design Thinking: Autoethnographic Insights and Design Implications. In Proceedings of the 2023 5th World Symposium on Software Engineering (pp. 55-60).
- Tholander, J., & Jonsson, M. (2023, July). Design ideation with ai-sketching, thinking and talking with Generative Machine Learning Models. In Proceedings of the 2023 ACM Designing Interactive Systems Conference (pp. 1930-1940).
- Niya Stojmenova, Rebecca Price. Exploring the Nuances of Designing (with/for) Artificial Intelligence. Design Issues [https://doi.org/10.1162/desi\\_a\\_00613](https://doi.org/10.1162/desi_a_00613)

### Assignment

*This is the most important part of the project brief because it will give a clear direction of what you are heading for. Formulate an assignment to yourself regarding what you expect to deliver as result at the end of your project. (1 sentence) As you graduate as an industrial design engineer, your assignment will start with a verb (Design/Investigate/Validate/Create), and you may use the green text format:*

To design a digital interface for human-LLM co-design and derive implications for the field of human-LLM interaction and design processes, by 1) identifying parameters that affect the human-LLM co-creative process 2) defining a human-LLM co-design process, 3) proposing an interface for human-LLM co-design, and 4) testing

*Then explain your project approach to carrying out your graduation project and what research and design methods you plan to use to generate your design solution (max 150 words)*

Week 1-14: To generate knowledge & implications for the interface, my process will involve the following elements that are situated in iterative cycle of undefined order, timespan and scope: Literature – Prototype – Validate – Interface implication & design. These terms are detailed below.

- Literature (Creativity theory (Design creativity, LLM creativity, Human-computer co-creativity), LLM agentic frameworks, design process (descriptive models, method), Human computer interaction)
- Prototype (Low-medium fidelity prototypes that aim to test assumptions extracted from literature and get feedback)
- Validate (Do tests with students/experts, gain feedback from experts on the interpretation of literature or to get recommendations on adjacent research)
- Interface design (Design the architecture of the interface based on the building blocks of creativity and human/LLM capabilities and make it visually easy to use)

Week 14-20: This final weeks are reserved for the proper documentation of the weeks that preceded, preparing both the report and presentation.

### Project planning and key moments

To make visible how you plan to spend your time, you must make a planning for the full project. You are advised to use a Gantt chart format to show the different phases of your project, deliverables you have in mind, meetings and in-between deadlines. Keep in mind that all activities should fit within the given run time of 100 working days. Your planning should include a **kick-off meeting**, **mid-term evaluation meeting**, **green light meeting** and **graduation ceremony**. Please indicate periods of part-time activities and/or periods of not spending time on your graduation project, if any (for instance because of holidays or parallel course activities).

Make sure to attach the full plan to this project brief.  
The four key moment dates must be filled in below

Kick off meeting	~ 21th of May
Mid-term evaluation	~ 8-13 July
Green light meeting	~ 23-27 Sept.
Graduation ceremony	~ 28 - ... Oct

In exceptional cases (part of) the Graduation Project may need to be scheduled part-time. Indicate here if such applies to your project

Part of project scheduled part-time	<input type="checkbox"/>
For how many project weeks	
Number of project days per week	

Comments:

### Motivation and personal ambitions

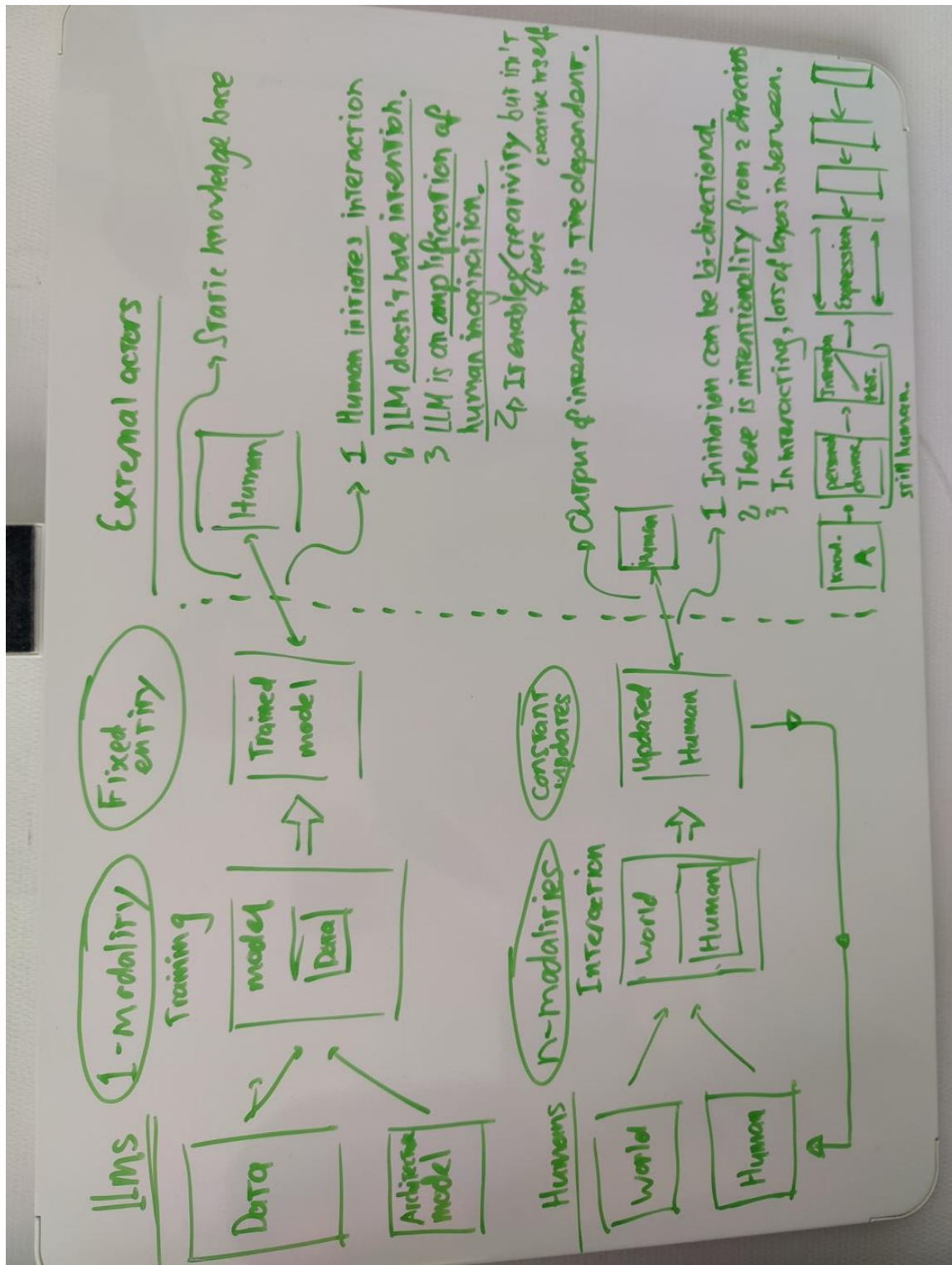
Explain why you wish to start this project, what competencies you want to prove or develop (e.g. competencies acquired in your MSc programme, electives, extra-curricular activities or other).

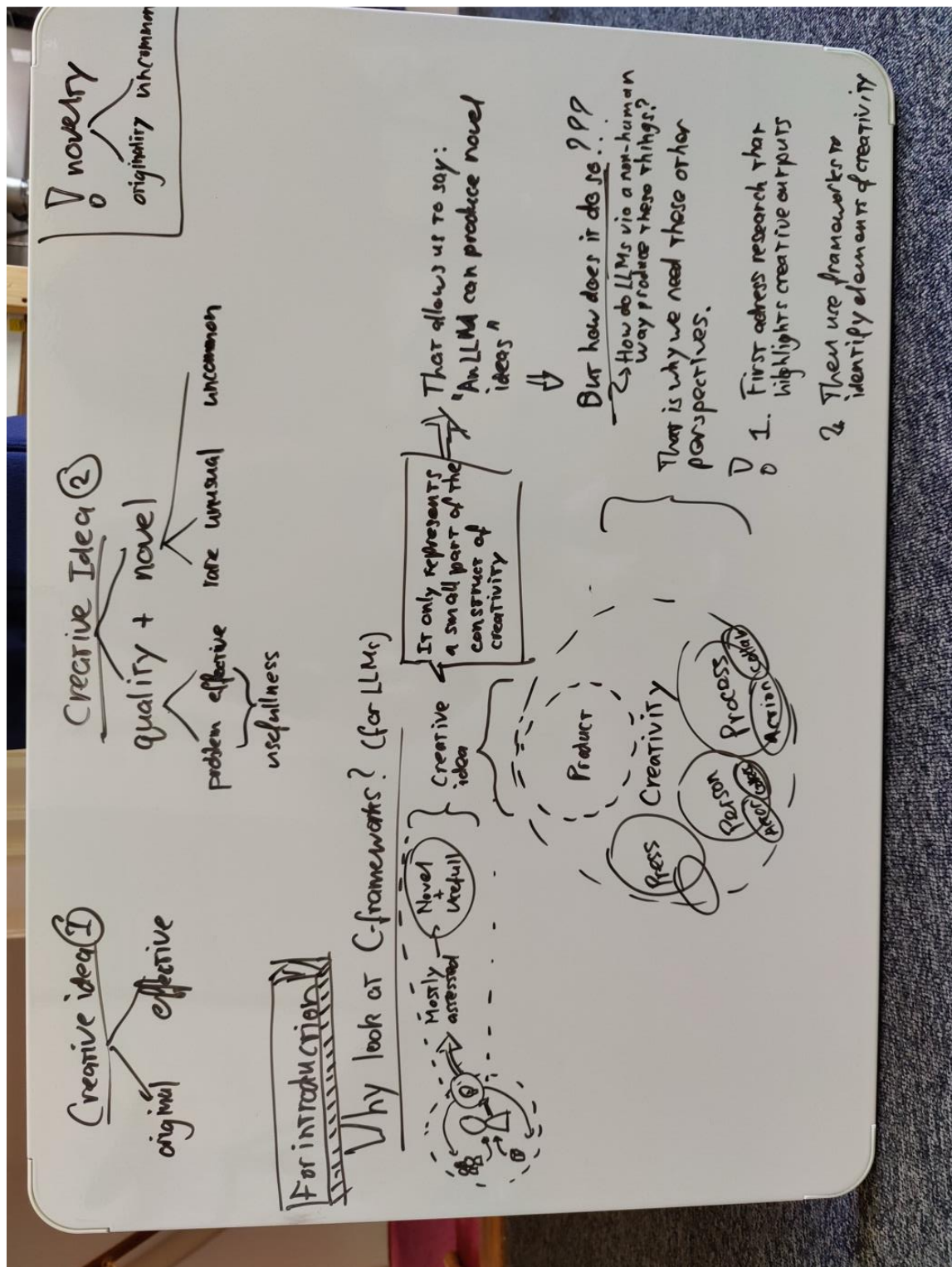
Optionally, describe whether you have some personal learning ambitions which you explicitly want to address in this project, on top of the learning objectives of the Graduation Project itself. You might think of e.g. acquiring in depth knowledge on a specific subject, broadening your competencies or experimenting with a specific tool or methodology. Personal learning ambitions are limited to a maximum number of five.  
(200 words max)

I wish to start this project as I am deeply interested in the future relation between human and computer intelligence and the place and value of both in our current society. I am currently working on a paper that evaluates LLMs in the light of current creativity theory to inform our vision on how to look at the opportunities and limitations of the current generation of LLMs in the domain of creativity. The aim of this graduation is to translate this theoretical investigation into a practical artifact, and to deepen my knowledge on the integration between design, creativity and LLMs. Finally, I am potentially interested in pursuing a PhD in this domain and am therefore very motivated to unravel this yet undefined domain of literature and identify threads for further research.

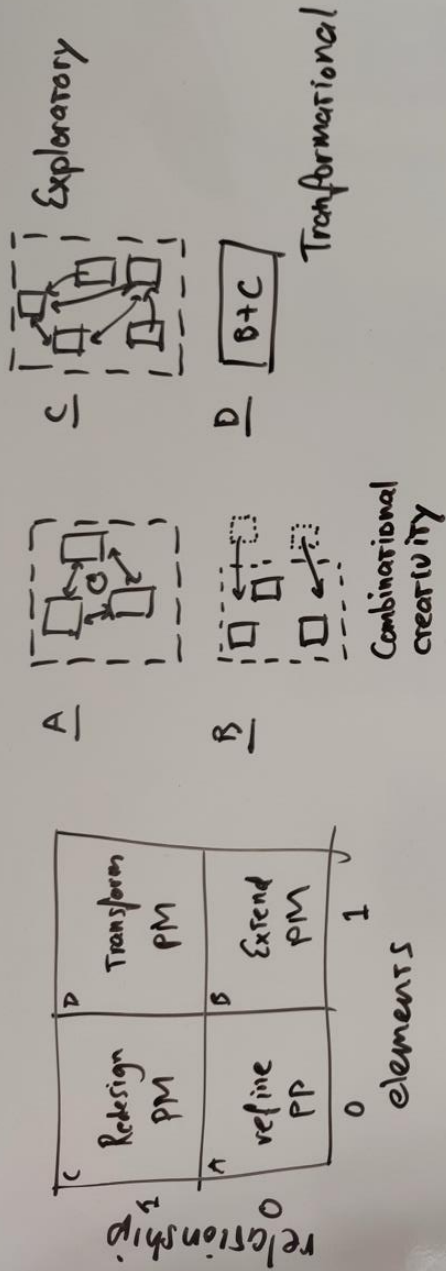


## Appendix B: Images of whiteboard drawings during literature review





Z



## Structure Paper

A

Move Baden  
to section  
"Landscape of  
creativity  
theory"

B

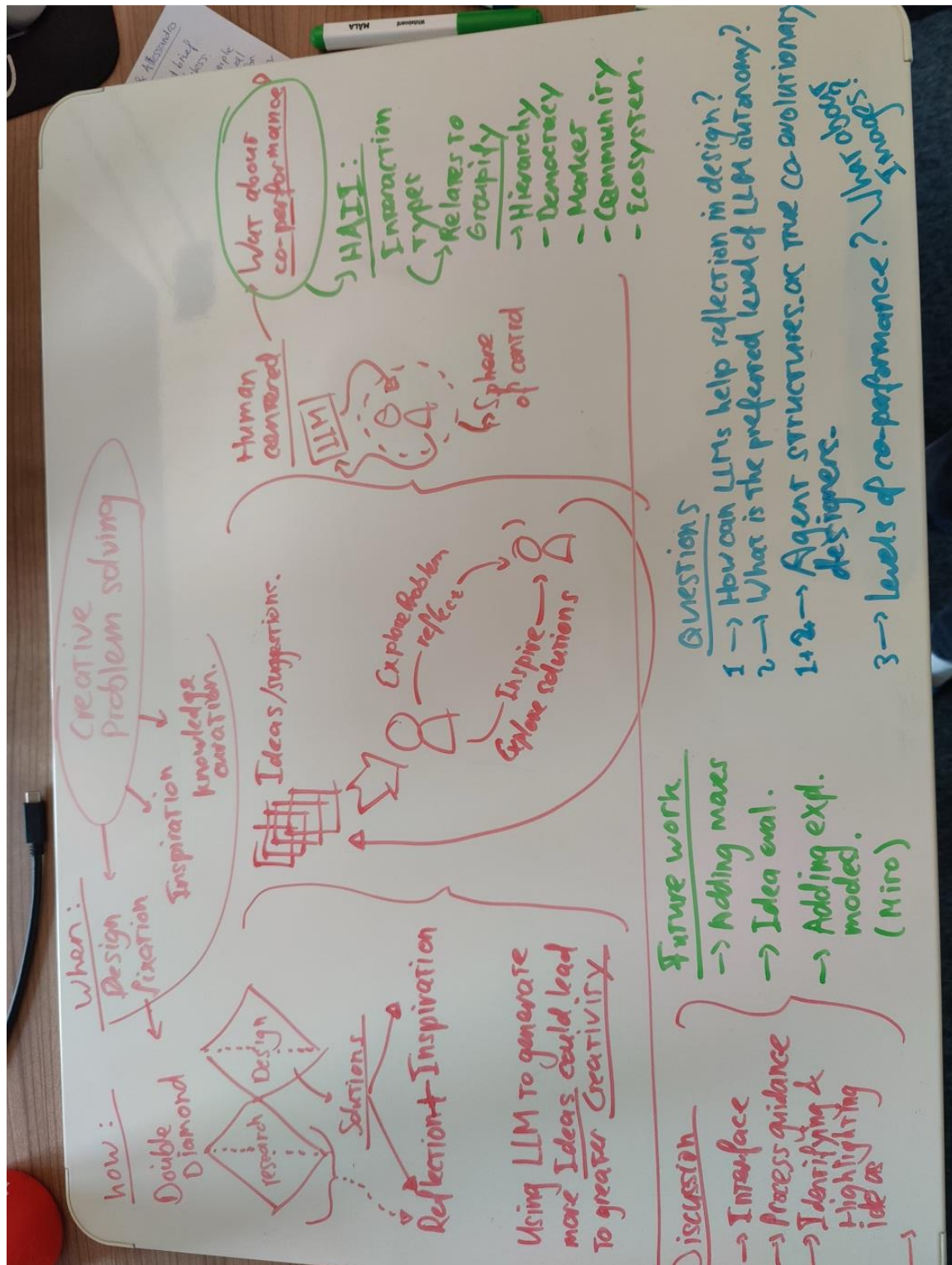
1) Impl. Baden  
into 'process'  
section  
2) Impl. creativity  
landscape into  
'product'  
section

C Instead of  $\downarrow$  &  $\uparrow$ , talk about  
Creativity theory on a micro & macro level

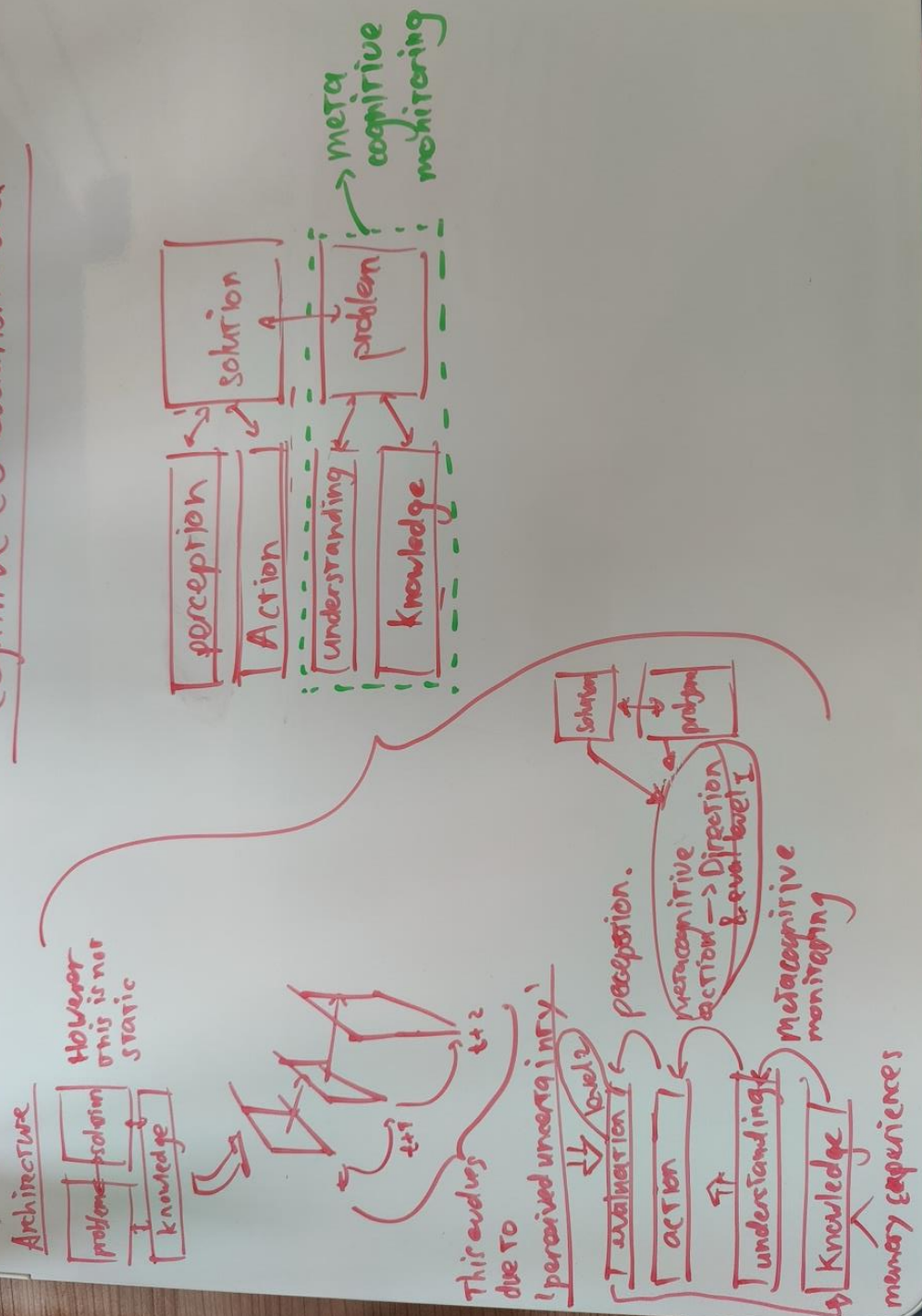
1) Micro: A research from several domains  
situated in one of the perspectives

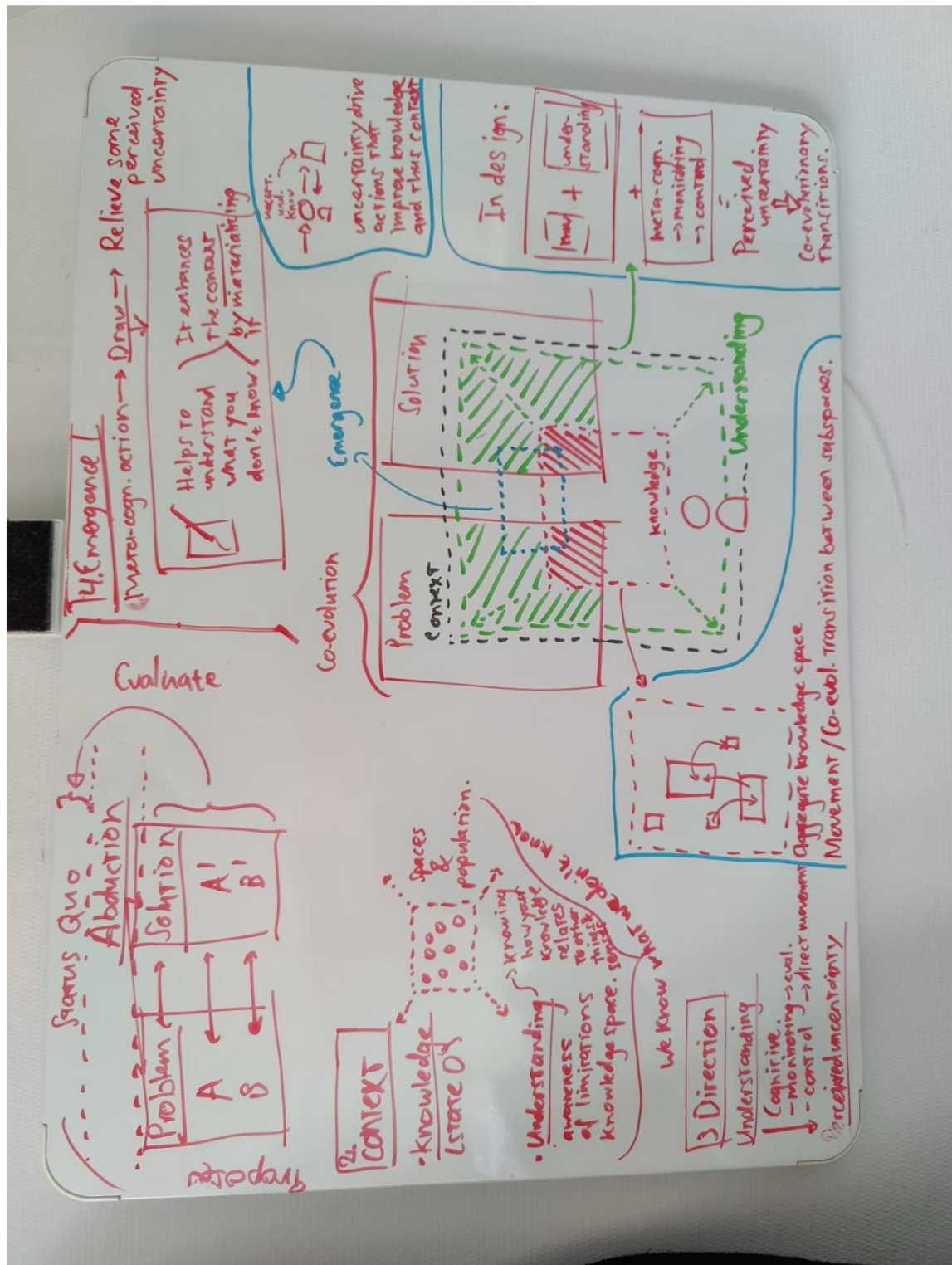
2) Macro: The creativity frameworks from  
both Psy. & CC



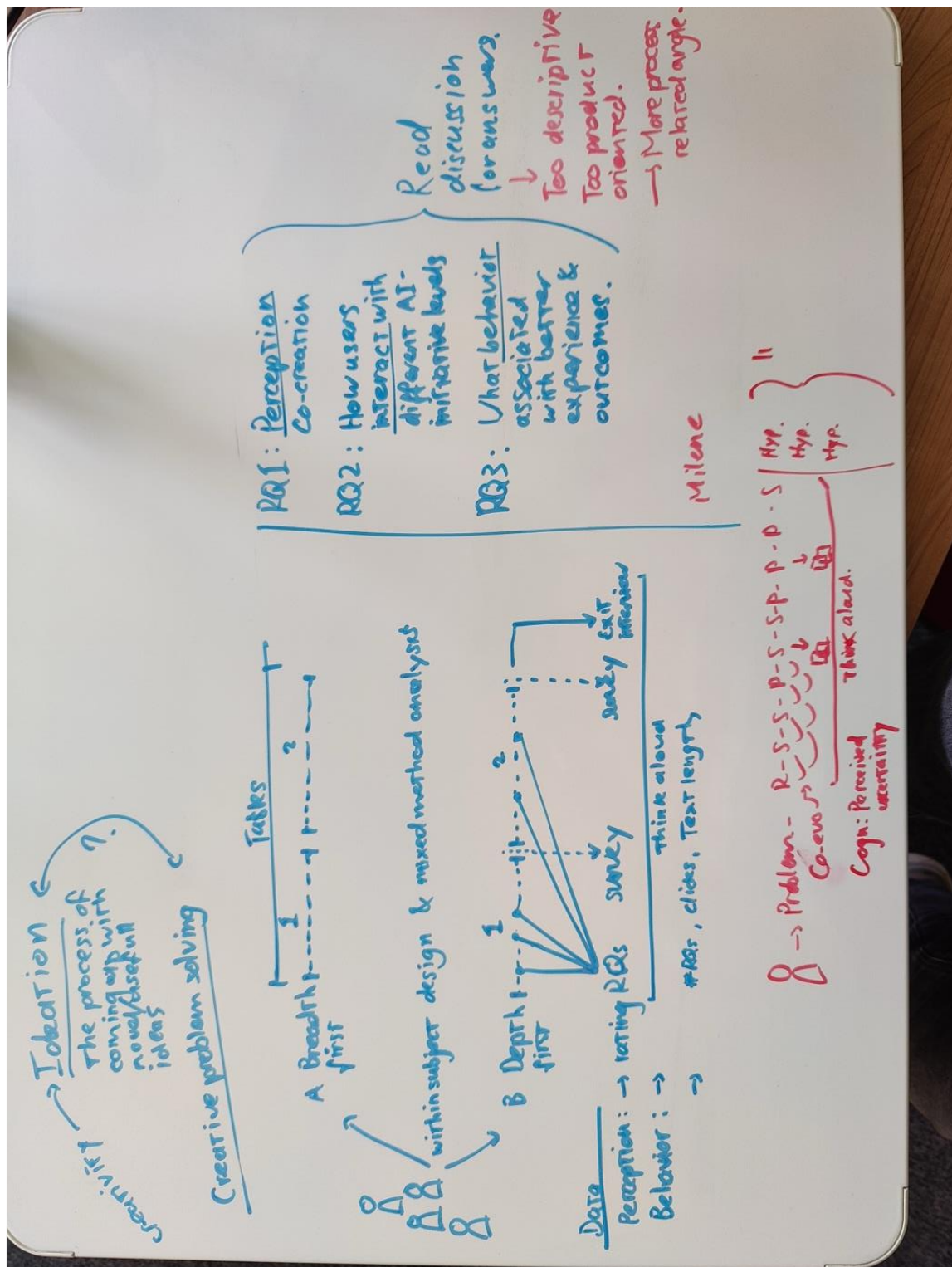


## Cognitive Co-evolution model



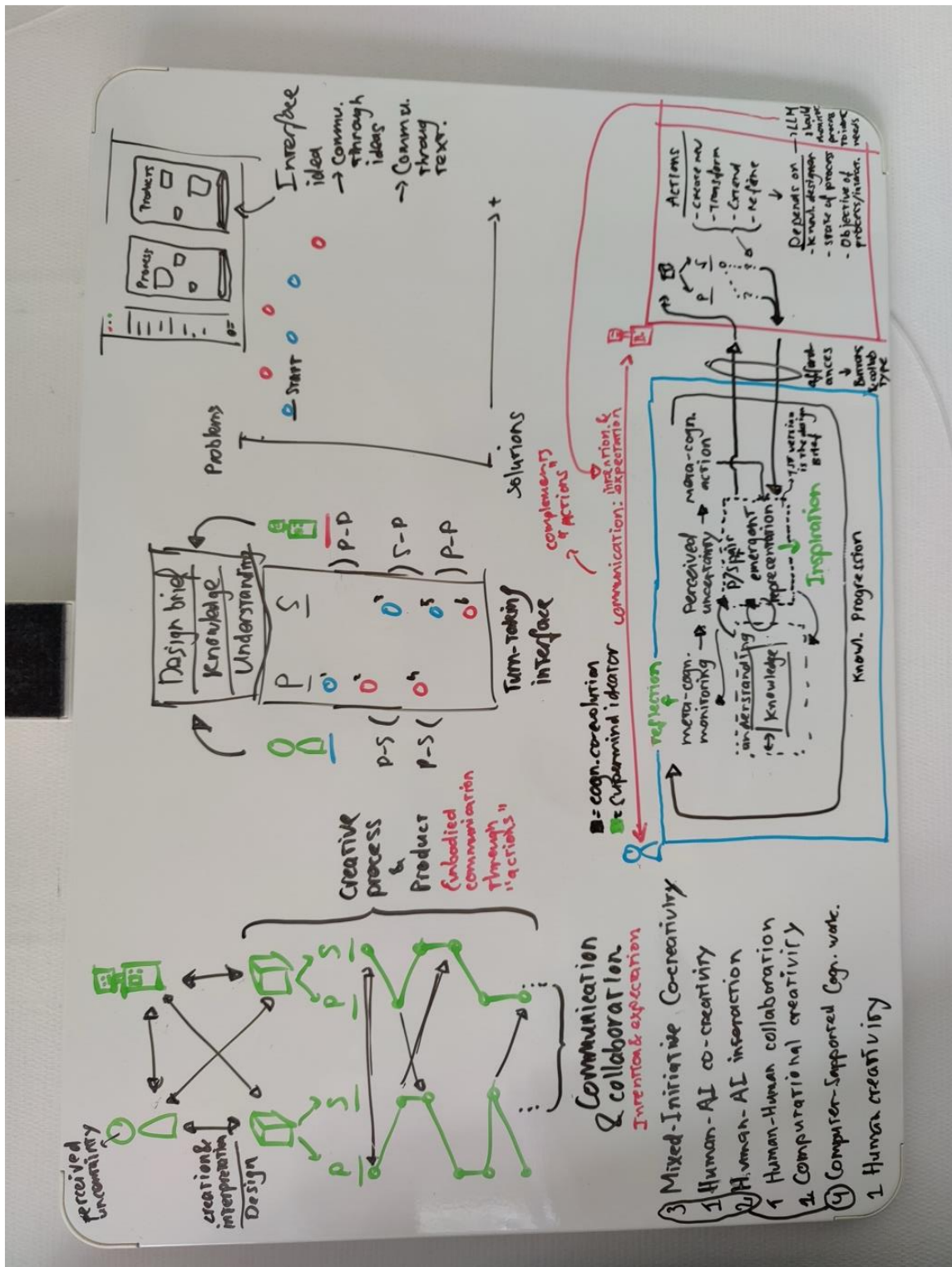


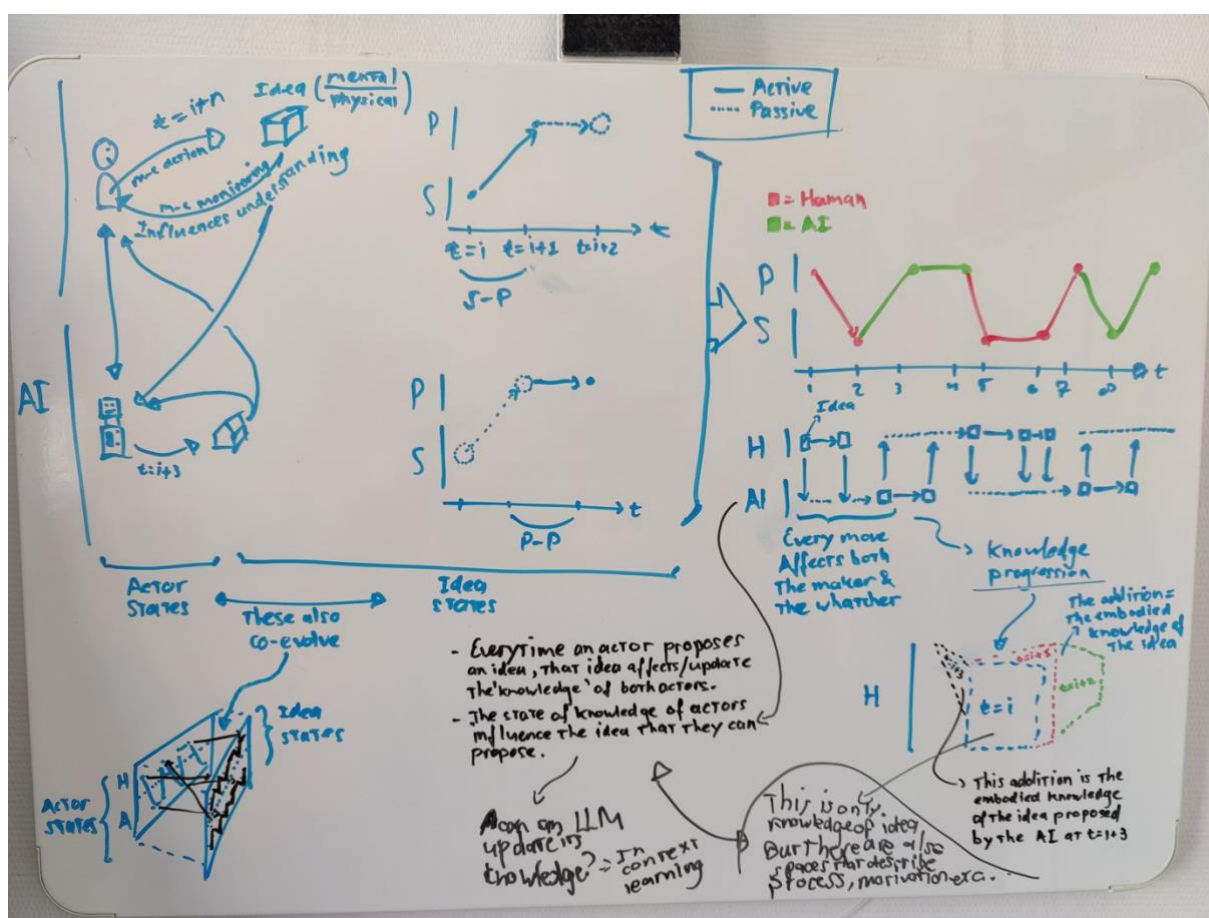


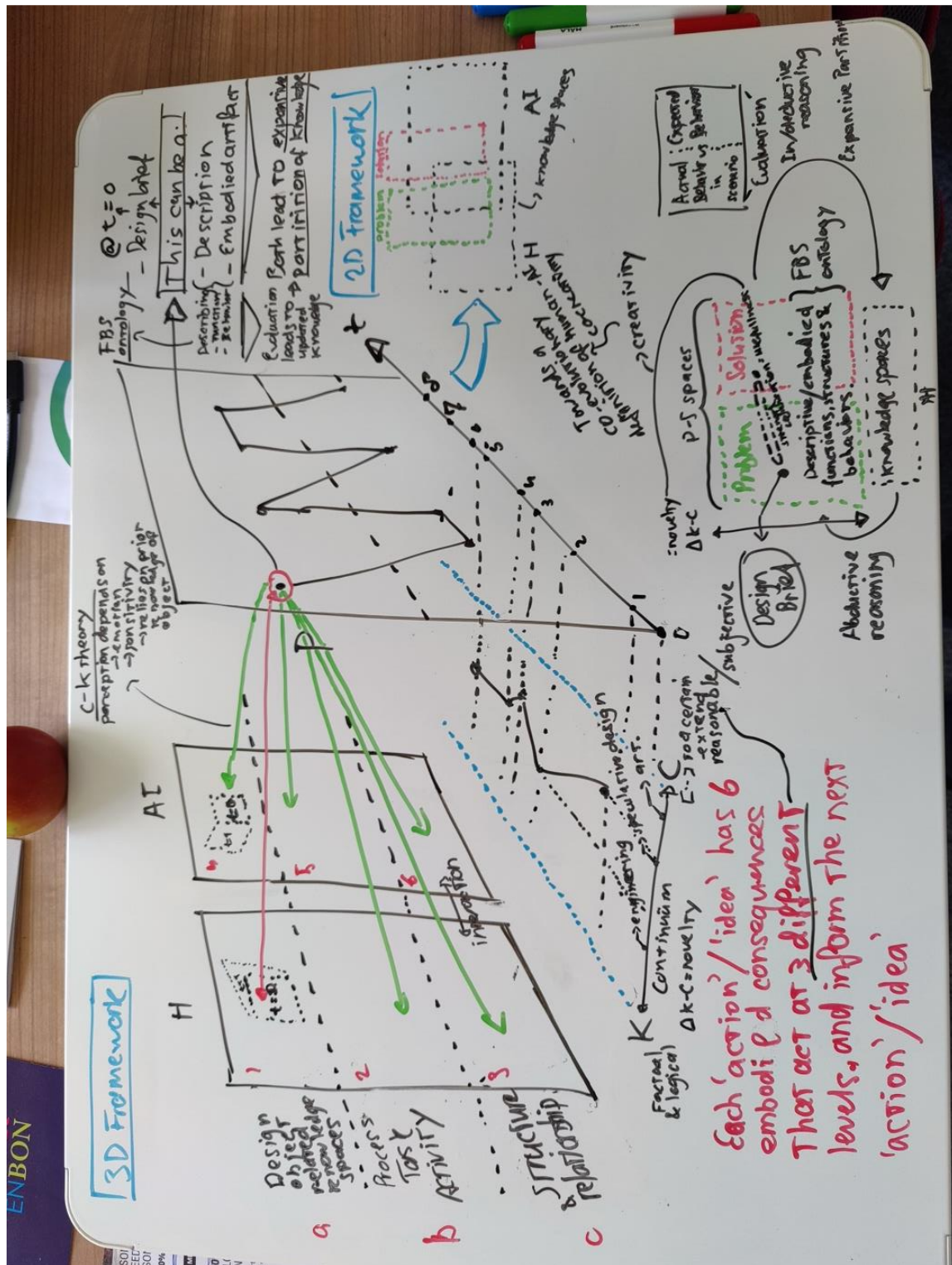








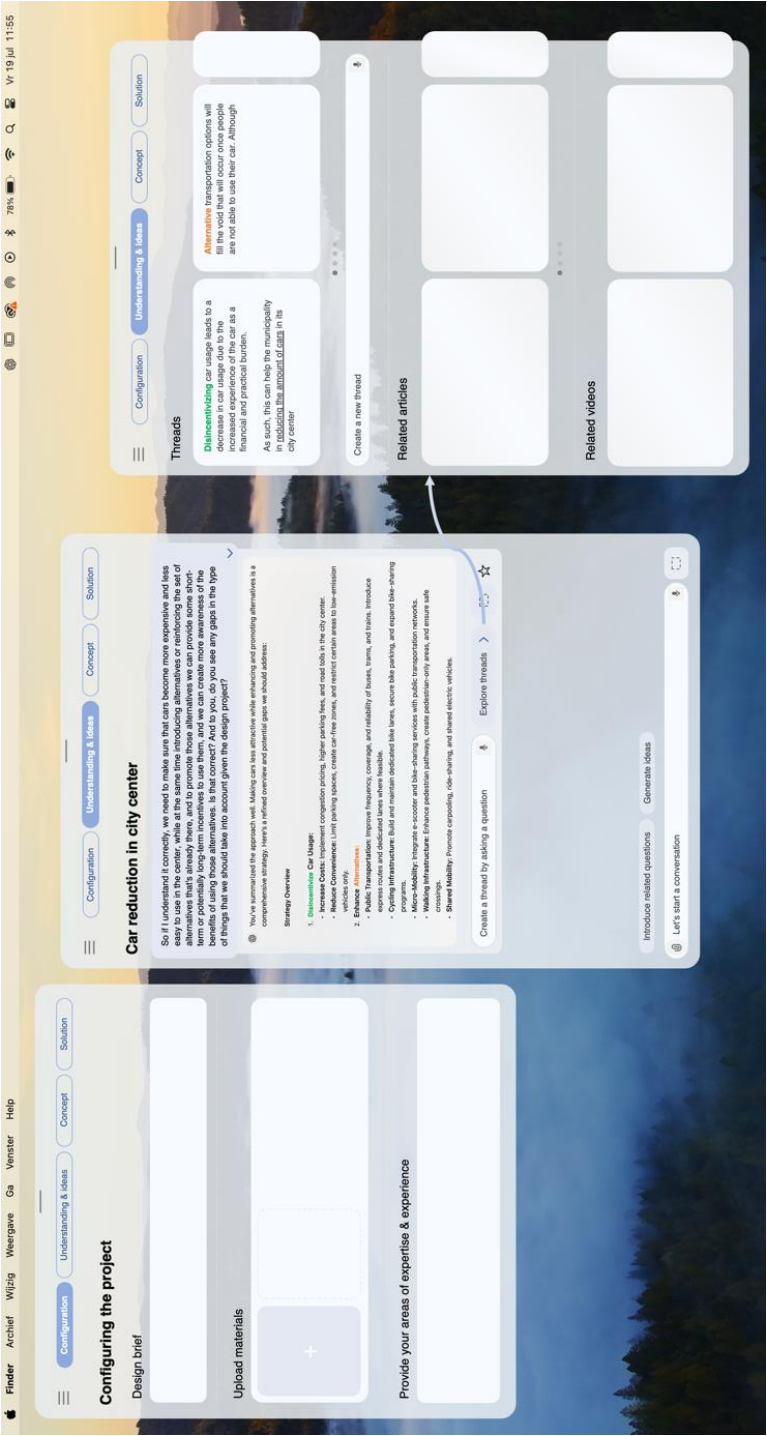






Appendix C: Augmentation system interface design iterations

Week 9



Week 12

