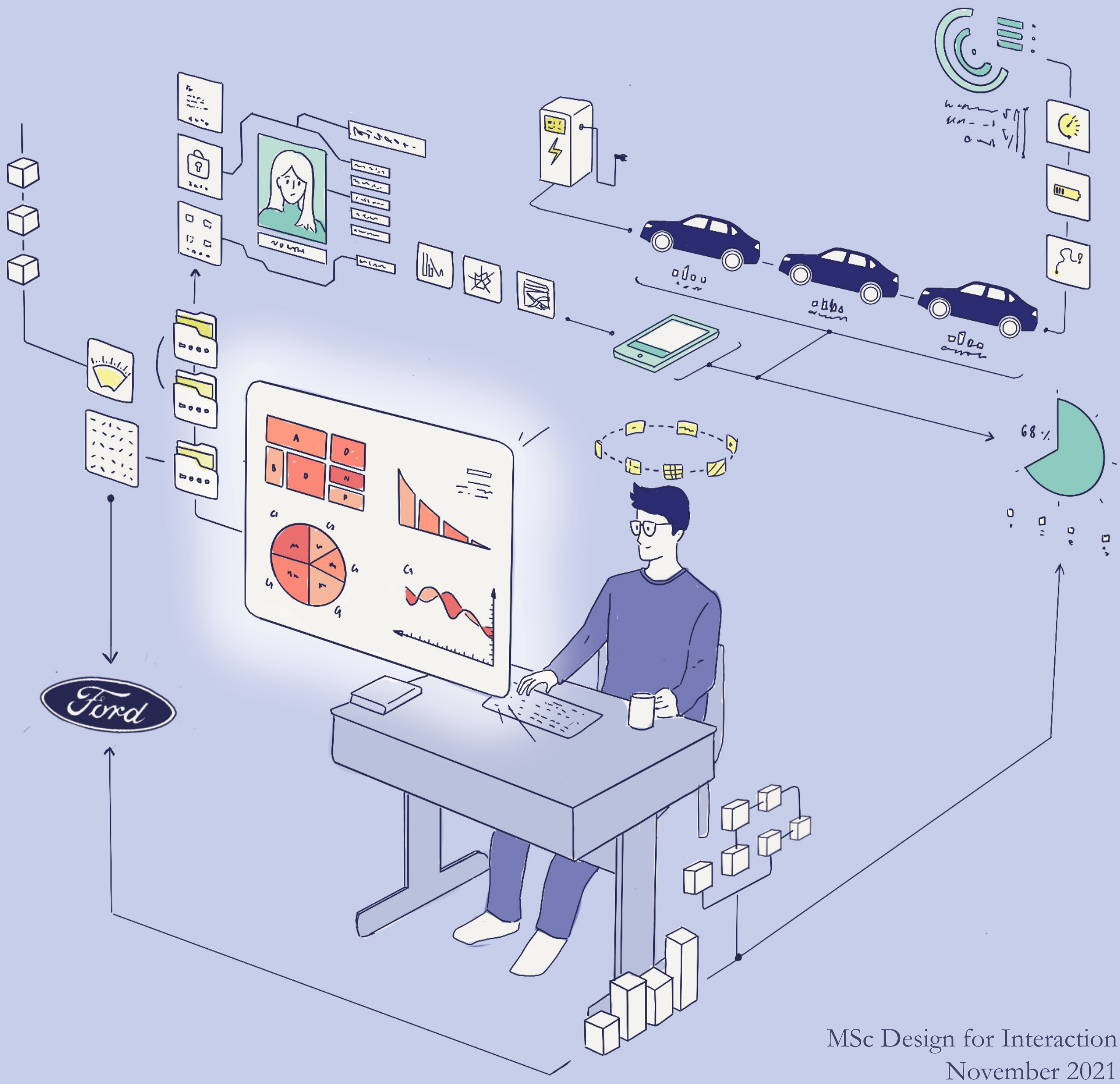
EXPLORATORY INQUIRING ENABLED BY DATA VISUALIZATION TO ENHANCE DESIGNER'S CREATIVE PROCESS

Déborah Mellado Cruz



MSc Design for Interaction
November 2021
Master Thesis

Exploratory inquiring enabled by data visualization to enhance designer's creative process.

Déborah Mellado Cruz
MSc Design for Interaction
November 2021
Master Thesis

Delft University of Technology
Industrial Design Engineering
Landbergstraat 15, 2628
CE Delft, the Netherlands
+31 (0)15 27 89807

Ford Research and Innovation Center
Süsterfeldstraße 200, 52072
Aachen, Germany
+49 241 94210

Supervisory team:
Dr. Milene Gonçalves
Dr. Senthil Chandrasegaran

Company supervisor:
Ir. Nicole Eikelenberg

“Everything is an oversimplification.
Reality is messy and complex.
The question is whether it is a useful simplification.
Know the limitations of an idea
and you can apply it to great effect
despite the messiness of reality.”

- James Clear

ACKNOWLEDGMENTS

Dear reader,

First, thanks for being here. Second, this section does not pretend to make total sense. I will write freely. I will be myself.

Finally, this project is ending, and these words are the last ones I am writing. Almost no time left to deliver. I wish I had more time to make a proper tribute to those who made this project possible. But, at the same time, every moment I tried to write this section, my own emotions totally overwhelmed me. Here I am crying in a company.

Are there even words that can express the gratitude feeling I am having? How much do I appreciate humans kindness and interconnection?

Forgive me for not having better words;

Thanks,

Thanks to my supervisors: Milene, Senthil and Nicole.

Thanks to my family: mum, dad, Alber, Ana and Tigre

Thanks Dami

Thanks Susana

Thanks to my girls: Cris, Inês and Maria

Thanks to my family in Delft: Francesco, Marco, Andrea, Edo and Francesca

Thanks to my ladies: Eva, Clau, Bea, Aldara, María and Soni

Thanks to my life coaches: Gonzalo, Ana, Carola, Ainoa and Guillermo

Thanks to my gang: Andrea, Javi, Felix, Ane, Beni, Jimmy, Sam, Carla and Simo

Thanks to my Ford girls: Tiara and Siqi

Thanks to my estrella friends

Thanks to all the lovely people who have helped me,

Deb

ABSTRACT

Ford aims to implement a data-enabled design approach, using data to inspire and inform the whole creative process. In this context, in 2020, they took a step to reach this goal by setting up a University Research Project with the Industrial Design Faculty at Technology University of Delft, proposing data visualisation as one of the research lines. In this master thesis project, the role of Data Visualization in Data-Enabled Design projects is investigated, specifically how to use Data Visualization more efficiently to generate insights and inform the creative process of Ford design.

The initial stages of the project were focused on understanding how Ford is using data in the design process and reviewing the published literature about creativity, data, data-enabled design, data visualisation, and exploratory data analysis. Then, through a critical reflection about the intersections of the theoretical research, eight possible design directions were identified in which data visualisation can support the creative process, termed "crossing bridges".

Afterwards, in collaboration with the company, the one with higher potential, "Exploratory inquiring", was selected to investigate further. Hence, three empirical studies were developed: personal explorations, a design students workshop and one workshop with Ford employees.

The last steps included the analysis and discussion of these sessions individually and comparatively, considering the limits of this research and the proposal of different aspects to consider in future research. Finally, I present a guide to support Ford in performing Exploratory Inquiring.

INDEX

1. INTRODUCTION 1

1.1. Project Introduction 2

1.2. Case study at Ford 4

1.3. Problem as given 5

1.4. Research question 6

1.5. Stakeholder Map 6

1.5.1. Ford 6

1.5.2. TU Delft 8

1.6. Design approach 9

1.7. General Overview 10

2. INSIDE FORD 13

2.1. Research Approach 14

2.2. Ford Data & Design Mindset 15

2.3. Departaments Involved 17

2.4. Accessible Data 20

2.5. Design Process & Data Use 24

2.6. Quantitative Data Process & Visualization 26

2.6.1. The use of data visualisation 29

3. THEORETICAL EXPLORATION 33

3.1. Research approach 34

3.2. Creativity 36

3.2.1. Definition of creativity 36

3.2.2. The creative process 37

3.2.3. Creativity, a constant play between solution and problem 38

3.2.4. The importance of knowledge in the designers' creative process 39

3.2.5. Questions in Design 40

3.3. Data Visualization 45
3.3.1. Definition of Data Visualization 45
3.3.2. The cognitive mechanisms to visually extract insights 46
3.3.3. The complexity of representation in Data Visualization 47
3.3.4. Data Visualization and Exploratory Data Analysis 48
3.4. The intersection: Crossing Bridges 50
3.4.1. Accessible and immediate insights 51
3.4.2. Visual Perception 51
3.4.3. Amplify cognition 51
3.4.4. Exploratory Inquiring 52
3.4.5. Communication efficiency 52
3.4.6. Inventive communication 52
3.4.7. Playful interactivity 53
3.4.8. Iterative parallelism 53

4. REDEFINING THE SCOPE AND CHOOSING THE DESIGN DIRECTION 55
4.1. Problem Statement 56
4.2. Design direction & Design goal 57

5. EXPLORATORY INQUIRING 59
5.1. Exploratory Inquiring 60

6. EMPIRICAL STUDIES 63
6.1. Overview of the studies 64
6.2. Analysis Method 65
6.3. Study 1: Personal Explorations 70
6.3.1. Objective 70
6.3.2. Setting 70

6.3.3. Procedure 70
6.3.4. Results 71
6.4. Study 2: Students Workshop 73
6.4.1. Objective 73
6.4.2. Research Approach 73
6.4.3. Setting 73
6.4.4. Procedure 75
6.4.5. Results 76
6.5. Study 3: Workshop with Ford 79
6.5.1. Objective 79
6.5.2. Research Approach 79
6.5.3. Setting 79
6.5.4. Procedure 79
6.5.5. Results 80
6.6. Discussion 84
6.6.1. Discussion Personal Explorations 84
6.6.2. Discussion Students Workshop 84
6.6.3. Discussion Ford Workshop 86
6.6.4. Comparison of the three experiences 87
6.7. Limitations 90

7. CONCLUSIONS **93**
7.1. Exploratory Inquiring Implementation 94
7.2. Future recommendations 96
7.1.1. Proposals for the limitations 96
7.1.2. Other points to consider in the future 96

8. REFERENCES **103**

APPENDIXES

I. Project brief
II. Interviews at Ford
III. Eris Taxonomy completed with personal examples
IV. 5W Problem Definition
V. Personal Explorations
VI. Students Workshop
VII. Workshop with Ford
VIII. Creative session

1 INTRODUCTION

This chapter's first section introduces the topic and its relevance. The second section focuses on familiarizing the reader with the user case at Ford and its general context. Subsequently, the third section presents the problem as given, followed by the leading research questions. The last section describes the steps of the design approach of this thesis and offers an overview of the activities.

1.1. Project Introduction _ *Page 2*

1.2. Case study at Ford _ *Page 4*

1.3. Problem as given _ *Page 5*

1.4. Research question _ *Page 6*

1.5. Stakeholder Map _ *Page 6*

1.6. Design approach _ *Page 9*

1.7. General Overview _ *Page 10*

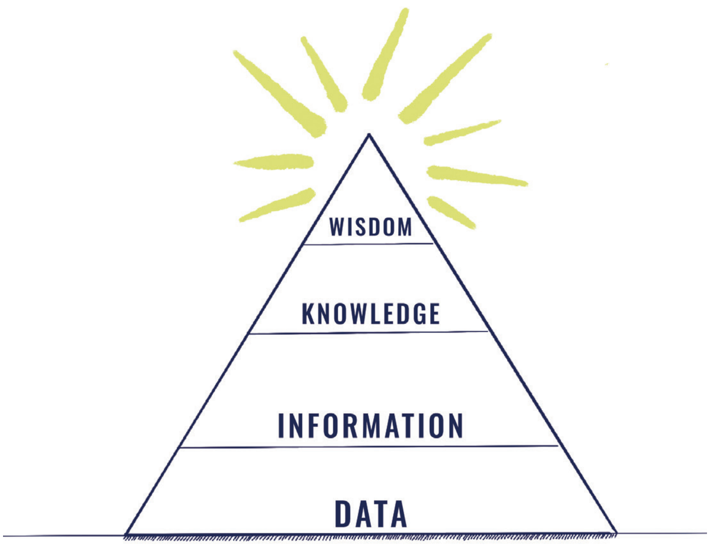
1.1. PROJECT INTRODUCTION

Data: "discrete, objective facts or observations, which are unorganised and unprocessed and therefore have no meaning or value because of a lack of context and interpretation" (Rowley, 2007)

After reading the data definition, it is hard to believe that "Design is and has always been informed by data" (King et al., 2017). The truth is that although data have "no meaning", it has the potential to be transformed into information, knowledge and wisdom (Ackoff, 1989) (Fig 1). And information serves the designer in multiple purposes like reducing the uncertainty in the design process; improving creativity; supporting awareness of previous solutions; developing an appropriate frame of reference for innovative design; enabling clear sharing and reception of knowledge within stakeholders, and facilitating and accelerating the idea generation process (Gonçalves, 2016).

Over the centuries, designers have refined ways of getting data to be leveraged into information. For example, they have mastered techniques like interviews, field observations, surveys or lab-based experiments to obtain data that allow them to understand the behaviours and needs of users. This kind of data is labelled as "thick" because it is highly contextualized and "enables the researcher to reflect upon how and why people do what they do" (Bornakke and Due, 2018).

Nowadays, designers are facing one of the fastest and biggest revolutions in data terms. Digital data is becoming omnipresent thanks to the increment of connected products and the constant use of digital services, resulting in high volumes of data arriving in different streams from millions of users. Besides, a lot of this emerging data has lower context complexity, so it is harder to get insights on human behaviour and has been denominated as "thin" (Bornakke and Due, 2018).



On the other hand, this shift in society is not just being driven by the growing abundance of data; it is fuelled by the development of technologies that change how we gather, store, analyse, and transform data like Artificial Intelligence computing or Cloud solutions.

Hence, all these innovations have tremendously increased the opportunities for designers to discover insights about users' contexts, habits, preferences or behaviours (Marti, Megens and Hummels, 2016). Nevertheless, at the same

Fig 1. The DIKW pyramid hierarchy is a well-known framework that helps to understand how raw data can result in useful insights. Adapted from Ackoff (1989).

time, is it a big challenge for designers to make use of this new world of data, as in general, they lack the knowledge and ability to use this emerging material in their activities (Davenport et al., 2019; Jung et al., 2019; Kun et al., 2020; Data-Centric Design-Lab, 2020).

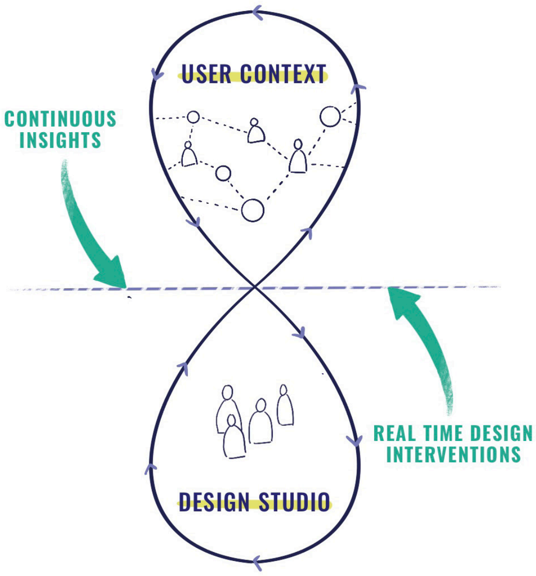


Fig 2. Data-enabled design approach: The designers' creative process is constantly fed by the data from the user context. Adapted from van Kollenburg and Bogers (2019).

In most established data-design approaches, this data is essentially used for evaluative purposes (e.g. data-driven, data-informed, data-aware design) (King et al., 2017) and in the latest stages of a design project. However, there is an urgent need to find ways to incorporate data into all the phases of the design process (Churchill, 2012; Kun, Mulder and Kortuem, 2018). Therefore, new lines of research aim to use data as a creative design material that can inspire and inform the design process from its beginning till its end. (Bogers et al., 2016) (Fig 2). These emerging design approaches are usually categorised as Data-enabled design.

One opportunity to bring data-enabled design approaches to reality is to use existing data-science tools and techniques with creative and exploratory purposes. In this context, a growing number of data-design researchers have highlighted the potential of data visualization as it can "empower" designers to identify and discover insights about products and users' behaviours (Gorvenko et al., 2020), support the understanding of live data streams (Wolff et al., 2016) or facilitate the contextualisation of big data on design projects (Knaflic, 2015). Furthermore, in the data science field, practitioners are convinced about the critical role of data visualization not only, as commonly used, at the end of a project to communicate results (Few, 2014) but in the early stages of research to explore and generate meaningful insights (Tukey, 1980) and to improve the speed and quality of the knowledge obtained (Batch and Elmqvist, 2017).

Despite the significant potential shown, the value that data visualization can bring to design remains unexplored (van Breemen, 2019). This project intends to inform and explore how data visualization can positively influence and support designers' creative process. Finally, this project aims to be used on a practical use case at the Ford automotive enterprise, which interest drove the motivation of this thesis.

1.2. CASE STUDY AT FORD

The automotive industry is not excluded from the data omnipresence. The amount of vehicle data is growing exponentially as a consequence of the "four technology-driven megatrends that are disrupting the industry – Autonomous driving, Connectivity, Electrification, and Shared mobility" (often referred to as ACES) (McKinsey & Company, 2019; PriceWaterhouseCoopers, 2020).

These fundamental changes offer new opportunities and challenges to automakers' companies like Ford, shifting from traditional vehicles to a myriad of mobility solutions, including digital services. A key point to innovation and to acquiring a significant competitive advantage inside the market is to optimise their use of data.

The importance of managing the vehicle data was estimated as an overall worldwide revenue pool of 450 - 750 billion USD by 2030 (McKinsey & Company, 2016).

Ford has already taken a place in the data race. In the crisis of 2008, the company profitably overcame their losses estimated at $17 billion by adopting the mindset that the CEO at that time, Alan Mulally, brought to the company: "Data will set you free". Since then, data analytics has slowly become a key to base all their decisions rather than anecdotal evidence, prejudices, or bias. By 2014, data-driven decision-making became the primary strategy throughout the entire enterprise (Henschen, 2017)

Ford's design departments were pioneers to understand the need to explore data to understand customer preferences and new user needs, according to Gartner's automotive analyst Thilo Koslowski (2012).

Ford's next ambition is to implement a data-enabled design methodology aiming to fully integrate data as a creative design material that can inspire and inform the design process (Ford proposal, 2020).

In this context, the Ford Research and Innovation Centre in Aachen (Germany) took a step to reach this goal by setting up a University Research Project, in 2020, with the Industrial Design Faculty at Technology University of Delft. Three lines of research were proposed, being one of them Data Visualization (Jansen, 2021). This proposal identified the potential of Data Visualization "to

"Like other automakers, Ford has been selling cars in the same way for over a century. However, as the market makes its likely shift to services, they need to figure out how customers will interact with those services and vehicles."

- Sam Abuelsamid (Ford Authority, 2017)

manage complexity and uncertainty, increase engagement and facilitate communication" and the opportunities that it offers "to better understand the user and its context and as an input for data-enabled innovations on products and services" (Ford Proposal, 2020).

1.3. PROBLEM AS GIVEN

The project's goal is to investigate the role of Data Visualisation in Data-Enabled Design projects in the context of Ford's Research Innovation Center Aachen. Specifically, explore how to use Data Visualization more efficiently to generate insights and inform the creative process of Ford design.[1]

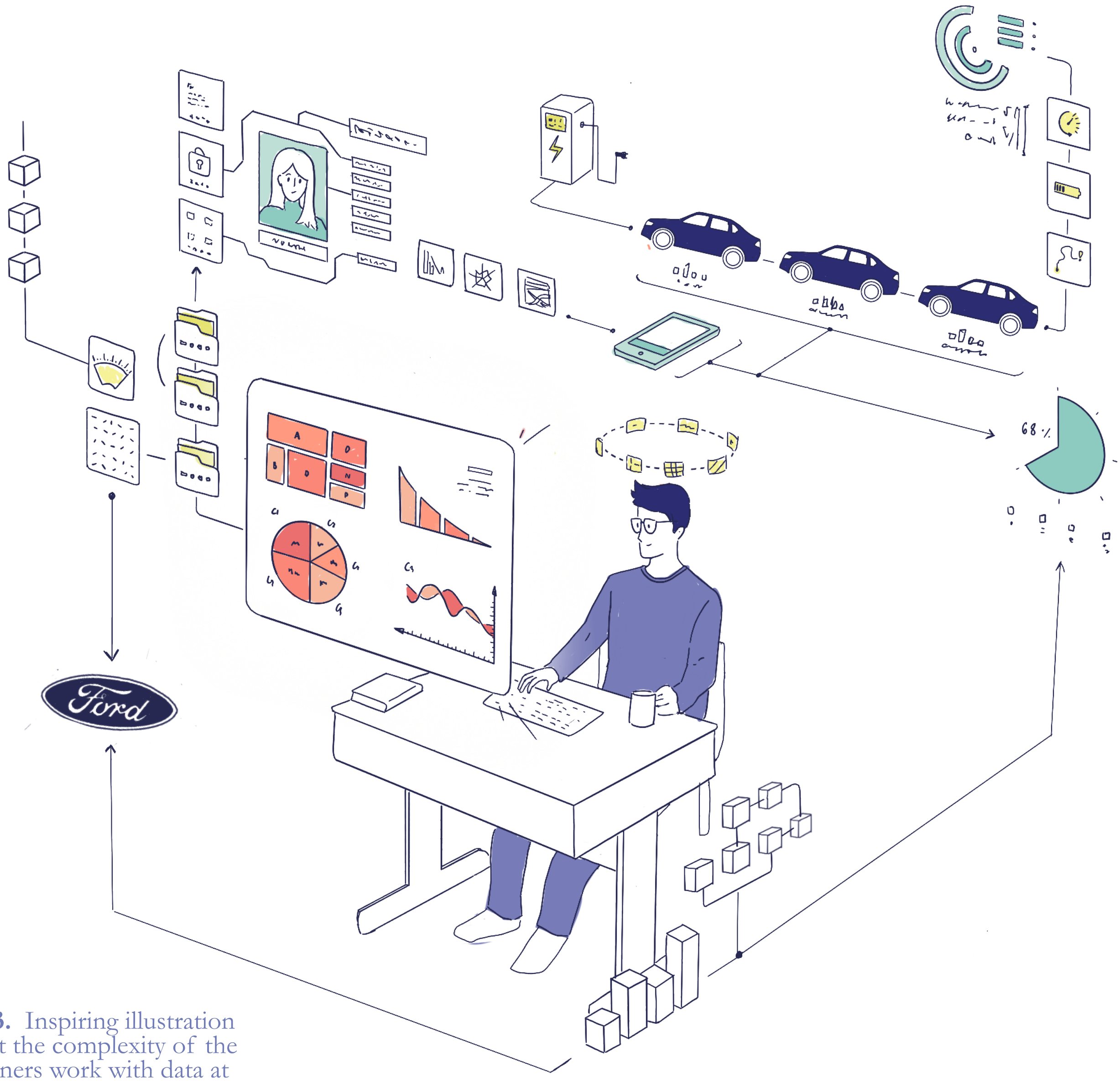


Fig 3. Inspiring illustration about the complexity of the designers work with data at Ford

1 The initial proposal is collected in Appendix: Project brief

1.4. RESEARCH QUESTION

The following research question led this project:

How can the ***Ford Smart Vehicles Concept*** team use ***Data Visualization*** as an exploration tool to facilitate their ***creative process***, leading to innovative insights?

1.5. STAKEHOLDER MAP

This project is part of the University Research Projects (URP) between the Industrial Design Faculty at Technology University of Delft (TU Delft) and the Ford Research and Innovation Center (RIC) from Aachen. This URP was established to find new opportunities for developing a data-enabled methodology from the initial phase of product and services creation.

The info-graphic in the next page displays the involved stakeholders and their relationships (Fig 4). The project is embedded in the mobility context, and its parties can be classified into two main groups: the company itself (Ford) and the university (TU Delft).

1.5.1. FORD

As mentioned previously, data analytic became a key in Ford's design process. In addition, Ford's ambition is to implement a data-enabled design methodology, so to delve into this knowledge is Ford's primary goal. In consequence, they collaborate with active involvement in this project. The departments implicated are mentioned below.

> The Ford Research and Innovation Center in Aachen (Germany)

Ford initially created the Research and Innovation Centers (RIC) to pursue technical advancement like new material applications. The Aachen centre, the only one in Europe, has expanded these activities and nowadays is also focused on innovation by creating new services or developing new working methodologies. To satisfy these goals, the

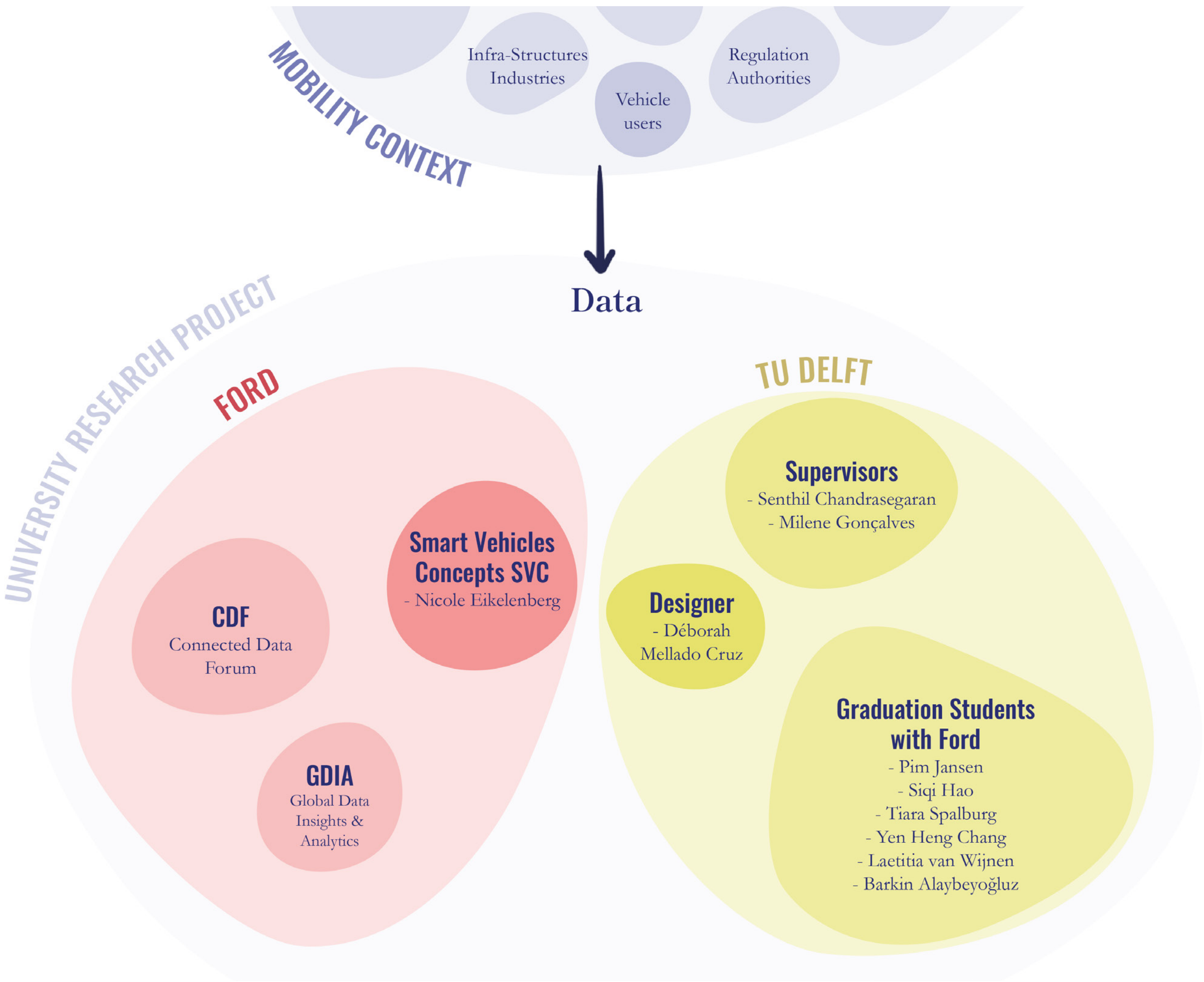


Fig 4. Stakeholders map of this project

multiple departments situated on Aachen collaborate with several universities and research institutes like RWTH Aachen, the KU Leuven, and the TU Delft.

» **The Smart Vehicles Concepts (SVC):**

This department is focused on developing innovative mobility concepts to be implemented in the medium and long term. Their projects follow a design thinking methodology and are grouped under various tracks and involve different employees.

» **The Global Data and Information (GDIA):**

This department offers a wide range of data analytic services inside Ford, supporting the project teams to develop connectivity and mobility solutions. Project teams usually approach the data specialists for data analysis requests.

» **The Connected Data Forum (CDF):**

A newly launched forum, constituted by representatives from different departments (mostly data specialists), aims to support everyone to get the correct data for individual use cases.

Ford's participation is led by **Nicole Eikelenberg - Company coach**, who has a fundamental role during the project, participating in collaborative sessions and bringing ideas to define and achieve the goal.

Multiple employees from RIC Aachen have been interviewed to get a good overview of their respective involvement and problems in their processes. An overview of the process of these interviews can be found in Appendix: Interviews at Ford.

1.5.2. TU DELFT

As a graduation project, the project will follow the TU Delft Faculty of Industrial Design Engineering guidelines. The main involved actors are:

> **Milene Gonçalves & Senthil Chandrasegaran - Chair and mentor**

Both supervisors have different expertise areas that complement each other for this project: Milene will help with the design process and creativity aspect while Senthil in Data visualization.

Milene Gonçalves is an assistant professor of Creativity in Product and Service Design at the Faculty of Industrial Design Engineering. As a specialist in creativity, creative problem solving, inspiration, design cognition, design methodology, design thinking and visual communication, she researches and teaches creativity, visual thinking and design processes. In addition, she supports designers to generate creative ideas by improving their use of inspiration and working in the co-evolution of problem and solution, framing and innovation.

Senthil is an assistant professor at the Faculty of Industrial Design Engineering. As a specialist in human-computer interaction, information visualisation, design theory and design methods, he researches and teaches the development of tools and methods that allow human and artificial intelligence to collaborate creatively. In addition, he coaches students in data visualization and design theory and methodology.

> **Déborah Mellado Cruz - Conceptual designer**

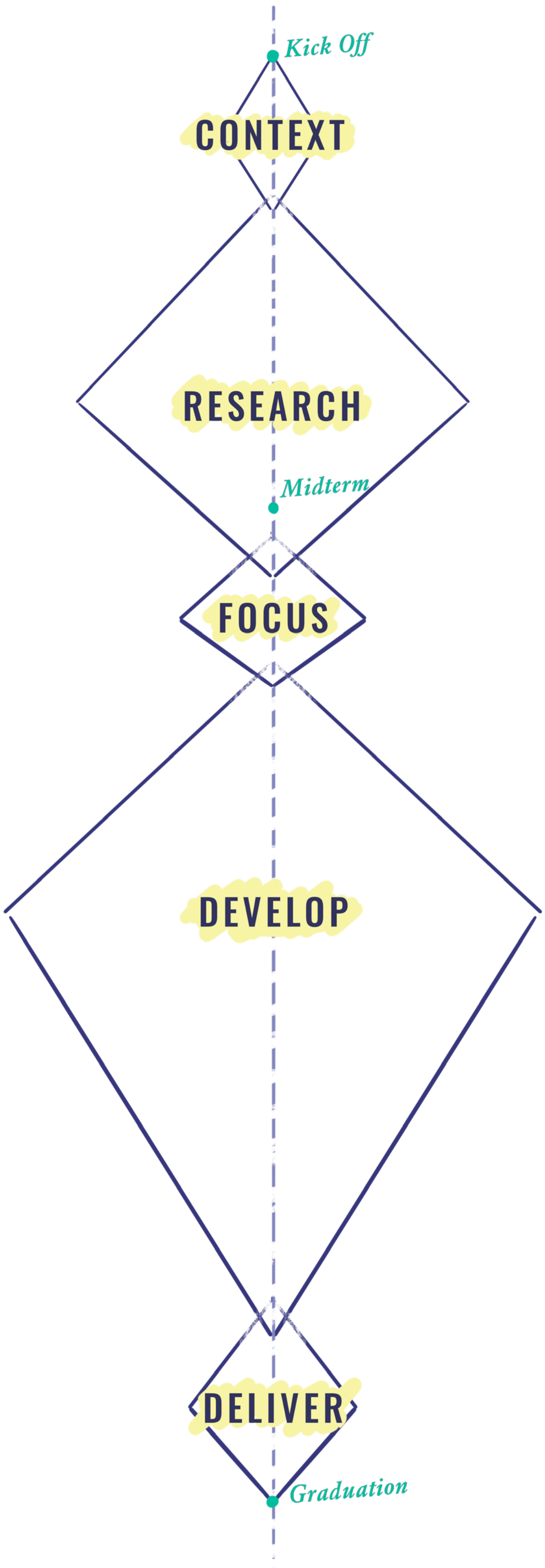
As the author of this thesis and Student of MSc Design for Interaction (TU Delft), my role is to investigate and deepen the knowledge of Data Visualization as a valuable creative tool for the Smart Vehicles Concept Design team (Ford) while maintaining a close collaboration with all stakeholders and ensuring an optimal result.

> **Ford graduation students**

Thanks to Ford URP, multiple students have collaborated in this project by amplifying our visions on data-enabled design.

1.6. DESIGN APPROACH

The structure of this project consisted of multiple iterations following a double-diamond process (Design Council, 2019), where convergent stages follow explorative and divergent phases. The following visual (Fig 5) correlates the different stages of this project and the chapters of this report:



CONTEXT

Initial research focused on understanding what is data and the actual relationship between data and design.

RESEARCH

User research about the current situation inside Ford. Literature review on the topics of creativity and data visualization.

2. Inside Ford *(Page 13)*
3. Theoretical Exploration *(Page 33)*

FOCUS

Structuring and converging the research insights to define the challenge scope. The final output of this phase is a clear problem statement and a specific design goal.

4. Redefining the scope and Choosing the design direction *(Page 55)*

DEVELOP

The design problem serve as a starting point for a double exploration which consisted of the development of a decision-making framework and empirical studies. The theoretical and practical dimensions build on each other; the results of the exploration in one dimension feed into and influence the exploration in the other dimension.

5. Exploratory Inquiring *(Page 59)*
6. Empirical Studies *(Page 63)*

DELIVER

Thanks to the development phase, the project converges on a proposed taxonomy, a final discussion and future recommendations.

7. Conclusions *(Page 93)*

Fig 5. The approach of this project followed convergent and divergent phases

1.7. GENERAL OVERVIEW

The following scheme gives a simplified overview of the activities conducted through the project.

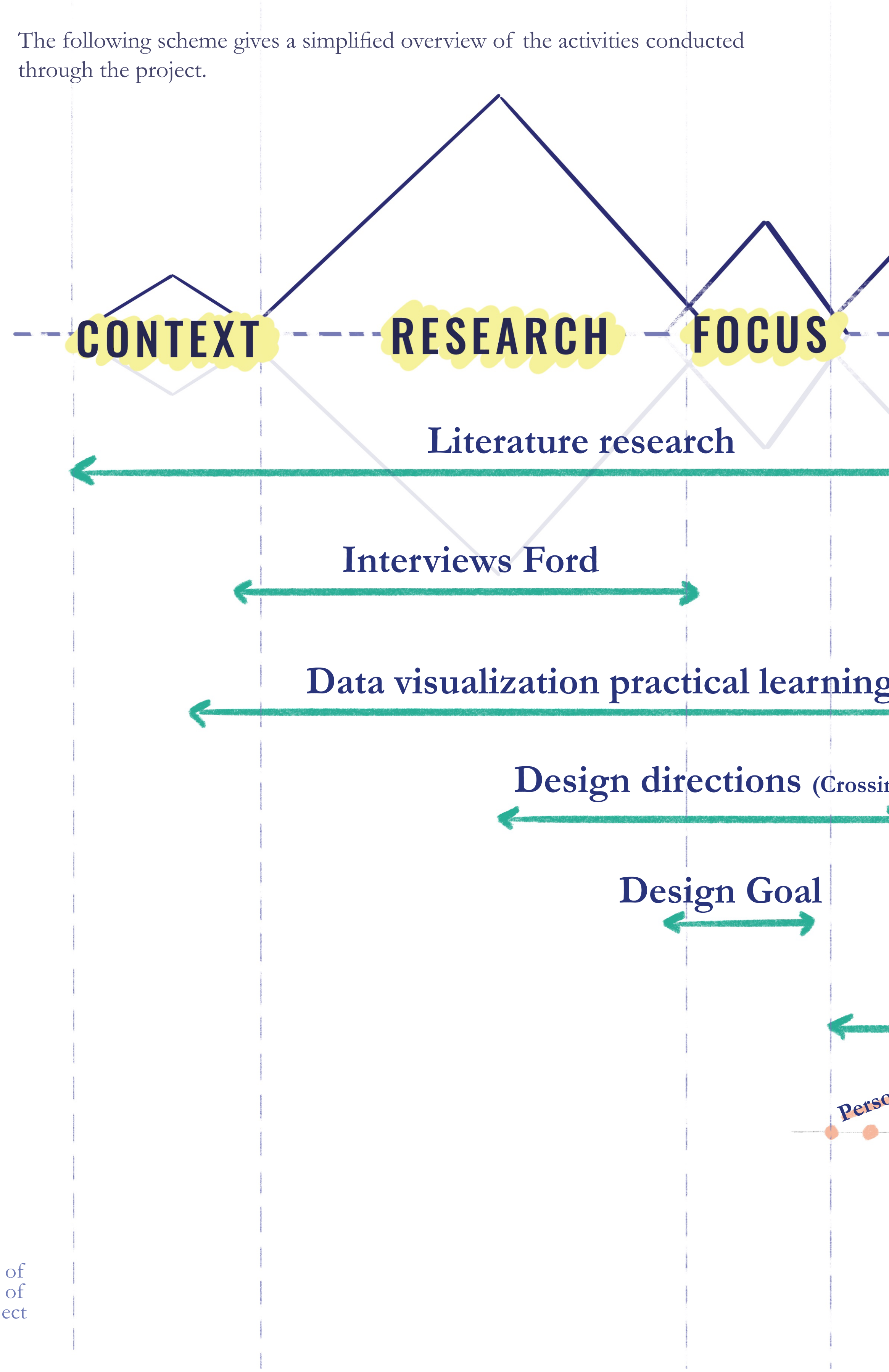
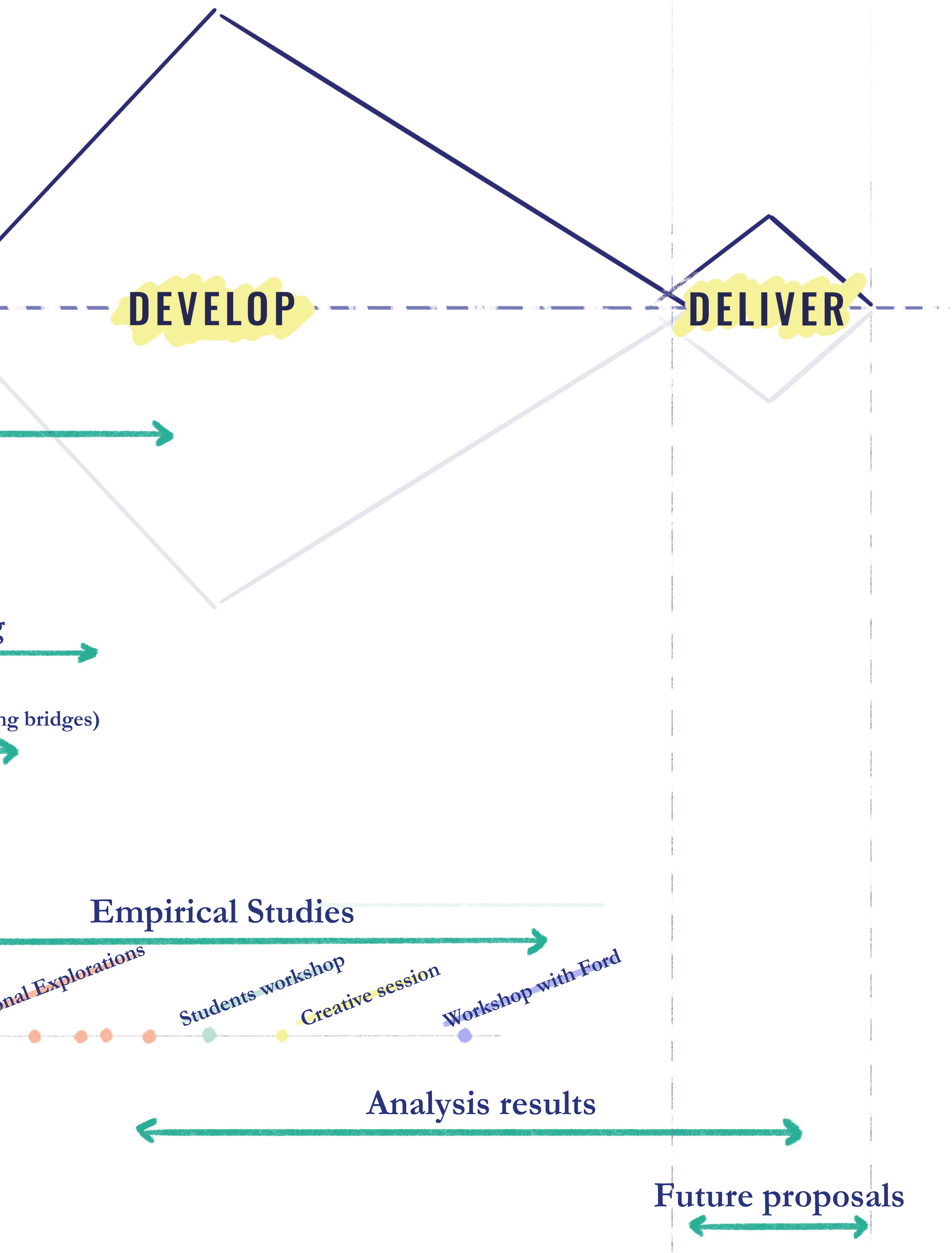


Fig 6. General overview of the activities and phases of this project



DEVELOP
DELIVER
g
ng bridges)
Empirical Studies
nal Explorations
Students workshop
Creative session
Workshop with Ford
Analysis results
Future proposals

2 INSIDE FORD

In this chapter, you can understand the current use of data in the design process of the SVC department. The first section introduces Ford's mindset regarding design and data. The subsequent sections delve into the specific situation of the SVC team from RIC Aachen: the departments they need to collaborate with to use data, the data sources available, the tools and processes currently used to analyse data, the latest role of data visualisation and how it is being implemented along its creative process.

2.1. Research Approach _ *Page 14*

2.2. Ford Data & Design Mindset _ *Page 15*

2.3. Departaments Involved _ *Page 17*

2.4. Accessible Data _ *Page 20*

2.5. Design Process & Data Use _ *Page 24*

2.6. Quantitative Data Process & Visualization _ *Page 26*

2.1. RESEARCH APPROACH

The approach chosen for this project consists of a combination of techniques to understand the current situation at Ford:

» Desk research with available online Ford sources and including the observations made by other URP students in their MSc Graduation Projects (Jansen, 2020; Hao, 2020; Spalburg, 2020)

» Informal conversations in weekly meetings with the company supervisor: Nicole Eikelenberg.

» Observation of as a participant in a creative workshop led by another URP student Laetitia van Wijnen (21st March 2021)

» Interviews with different Ford employees, selected either based on expertise on the data or the design thinking process and, of course, availability (Appendix: Interviews at Ford) . Fig 7 offers an overview of the different interviewees and the departments involved.

SMART VEHICLE CONCEPTS TEAM

Design engineer
Focused on the implementation of Design Thinking and innovation.

Research engineer
Focused on the implementation of data from a design perspective.

User experience expert
Responsible of the user research of design projects.

Research engineer
Facilitating creative sessions.

GLOBAL DATA INSIGHTS & ANALYTICS

Data analysis expert
Data analysis and assisting the design team regarding data visualisation.

Data analysis expert
SyncAnalytics specialist

CONNECTED DATA FORUM

Data coach
Specialist on guiding the data process

Duration: 60 minutes
Setting: Webex or Zoom
Procedure: Semi-structured with questions prepared in collaboration with the other students. Roles were divided between students to be the interviewer or observer.

Fig 7. Overview of the different Ford's employees interviewed through this project

2.2. FORD DATA & DESIGN MINDSET

Ford Motor Company, from now on referred to as Ford, is a leading international company in the automotive sector, present in more than 125 countries worldwide, founded by Henry Ford in 1903.

Ford's story is intrinsically related to innovation, looking not only for new technology and product developments but also for new ways of working. In recent years, the automaker firm had experienced a crucial shift of mindset relevant to this project thanks to two principles:

> ### Data-driven:

As stated in the introduction, in recent years, a new mindset has been driven at Ford, adopting an operational strategy based on data analytics and implementing a data-enabled design attitude at all levels.

Ford approaches this stage being fully aware of its responsibility for the safe handling of data, adopts responsible practices and maintains a serious commitment to protecting the privacy of its customers while exploring innovative solutions.

"Ford's target is to become the most trusted company in the world". (Interviewee Connected Data Forum, 2021)

> ### Design Thinking and human approach:

In 2017, with the addition of Jim Hackett as CEO of Ford, a radically different approach was introduced to Ford's mindset: a human-centred design thinking approach. Design thinking is considered a process for creative problem solving (IDEO, 2020).

To ensure the change of mindset, Ford developed in collaboration with IDEO multiple tools and an educational program to transmit the human-centred design thinking values to all its employees. One of the main additions is a standardized model for the design process, which will be described further in the practical context of the Smart Vehicle Concepts team in Aachen (on section 2.5)

This new approach led to a complete organizational change focused on human services development. Multidisciplinary engineering, design,

"We are creating a future where things like mobility, autonomy, and connectivity are flexible, memorable, mindful, and most of all, empowering"
- D Ford

purchasing, and other teams strive to find out what needs to change to meet the constantly evolving customer needs. As a result, Ford is implementing a unique creative culture based on design.

To be human-centred does not just mean to put humans at the centre of the design problem, but to put humans at the centre of the design solution and empower them in different ways. (IDEO & Ford, 2019)

Implementing these two principles at all levels of the company is a great challenge. Consequently, employees need more support to truly incorporate design thinking and a data approach into their work activities. That is one of the reasons why RIC Aachen has decided to collaborate with TU Delft to learn more about integrating data into its design process.

- The company promotes the use of Data and Design as the core of all its operations
- Employees are encouraged to acquire a Design thinking mindset.
- Ford's environment seems perfect to shift from Data & Design theory to practice and discover new challenges & opportunities within the use of data in the design process.

2.3. DEPARTAMENTS INVOLVED

As presented in the section 1.5 , this project is focused on supporting the Smart Vehicles Concept Team located in the Ford RIC Aachen. To simplify the terminology, we will refer to it as the "design" team.

Supervised by Walter Pijls, this team fosters innovative human-centred vehicle transport solutions with a medium–long-term impact. Simultaneously, multiple projects are developed to obtain different outcomes:

» Explorative: its objective is to identify opportunities for new products or services.

» Discovery: carried out to determine the feasibility of a specific idea.

» Technology Development: focused on improving technical aspects such as materials or new technologies used for the vehicle's production.

» Product Development: directed to create a product or improve existing ones.

From a data perspective, Explorative projects are the most challenging for the department, so my priority is to support the team by focusing on them and trusting that the solution will be more easily scalable to the rest of the projects. In the table below, I summarize the challenges and opportunities on which I based my decision.

REASONS TO FOCUS ON EXPLORATIVE PROJECTS	
CHALLENGES	OPPORTUNITIES
Characterized by complex open problems with no clear goals and many possible directions	These kinds of projects can benefit from Data Enabled Design's main ambition to use data as "an enabler for design exploration" (Kollenburg & Bogers, 2019).
Almost no experience in implementing quantitative data in comparison with the other projects	The lack of knowledge and understanding of using quantitative data can offer a bigger chance to find the best way to gather information and solve an issue.
The lack of definition, the variability and dynamism of this project's type is very distant from the data scientist approaches (Kollenburg & Bogers, 2019). Therefore, they do not always get high priority from the GDIA or CDF.	If the team understands better how to implement data, it could be easier to involve the data analysts and get more appropriate resources.

Table 1. Main reasons why I believe explorative projects at Ford can benefit from a data-enabled design perspective

Each explorative project is usually led by a designer supported by a small team of employees with different backgrounds like design, engineering or marketing. Ford's design thinking mindset is present, placing their customers' needs as the heart of their process, although some employees are still not familiarized with design thinking techniques. It is essential to consider these differences in attitude towards this design working method to create a suitable solution for the team.

Thanks to the variety of expertise and Ford's data-driven mindset, an employee usually can handle small quantitative data sets. However, most of the time, to access the different vehicles data sources (check 2.4), they will need to collaborate with two departments:

» The Global Data and Information (GDIA) department: formed in 2015 and located partly also in Aachen, this department supports the Smart Vehicle Team with the data collection and analysis. Both teams have pointed out the need to improve their communication, especially after working online during Covid. One of the challenges is that the data specialist usually follows a deductive approach, reasoning from one hypothesis to reach a logical conclusion. At the same time, designers predominantly work with abductive thinking, which defines in parallel both the problem and the solution (also observed by Jansen, 2020). In other words, the GDIA is experienced in using data for optimization, while designers usually prefer to use data for exploration. Therefore, requesting GDIA for data with optimization purposes requires less effort than exploration, so the latter is usually not a high priority and are allocated fewer resources.

» The Connected Data Forum (CDF): To facilitate Ford's data-driven mindset, this group of employees has specialized in the data flow within Ford. The CDF aims to support any employee "in getting the right data for the individual use case" (CDF Interviewee, 2021). Mainly, they guide and advise through the whole data process, helping the Smart Vehicles to define an "appropriate" request for the data specialist. Currently, this team does not correspond to a single department. Nevertheless, it will become an official team under a supervisor and a manager with the company's transition towards full data connectivity. The background of the employees of this team is engineering, which makes it easier to empathize with the Ford employees who requested the data, often engineers in other departments.

The interactions between the three concerning the data will be tackled in more detail in section 2.6.

- Employees from the Smart Vehicle Concepts team have different backgrounds and different levels of knowledge in data and design thinking techniques.
- This thesis will focus on Ford explorative projects that tackle ill-defined problems and look for innovative opportunities requiring exploration and creativity.
- The SVC team has to collaborate with the GDIA and CDF to use data sources related to Ford vehicles. Therefore, the better the communication between them, the easier it will be for the SVC team to improve their use of data in the design process.
- The recent creation of the Connected Data Forum exemplifies a high involvement of the company to implement data.
- As GDIA is not used to working with data in a more exploratory way and designers also do not know how to do it, the SVC team struggles to get more resources and involvement from the data analysts.

2.4. ACCESSIBLE DATA

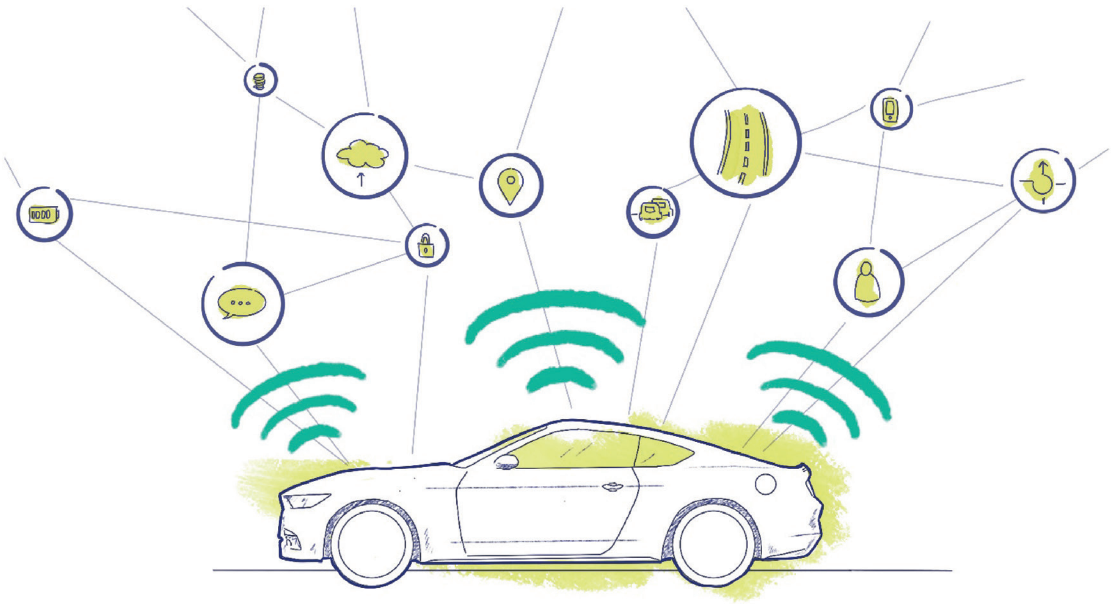
The Smart Vehicle team can make use of different sources of data. Therefore, the following classification was created based on the collecting sources, challenges, and possibilities (Fig 8 offers an overview of this classification).

> ### Data collected by the Smart Vehicle Team:

» #### Qualitative

Data obtained in the early research phase through interviews, observations, surveys, or other methods. Up to now, this data is the most commonly used by the team to understand the users' problems and context.

- Challenges: Designers face the challenge of trusting if what their participants "say" matches what they "do" in reality.
- Possibilities: To obtain this data, the team needs to recruit and involve participants. These can also be beneficial in later stages to discuss other data or co-create with them.



DATA COLLECTED BY SVC

Qualitative
Interviews
Surveys
...

Quantitative
PID
User test
...

VEHICLE RELATED DATA

Vehicle itself
Diagnostics
Driving data
...

Services
Ford Sync
Ford Pass
...

Other departments
Ford Customer Service
Vehicle Production
...

OPEN DATA

Related to mobility
Weather Forecast
Social Media
...

Fig 8. Classification of Ford accessible data sources based on my personal research

» Quantitative

Data acquired thanks to project-specific sensors installed in the vehicles, like Plug-In Devices (PID), to measure engine-related or driving parameters.

- Challenges: Designers need to recruit the participants and agree with them to install the devices.

- Possibilities: These sensors allow a higher control on the data collection and result in a limited volume of data. Therefore, it could be analysed by the SVC and may not need high resources from the GDIA. Additionally, also, in this case, they can contact back the participants for further research.

> Open data:

The team uses existing data from social media or other open sources, like observations from YouTube videos, Twitter threads, or Quora opinions between users. Nowadays, these ethnographic studies are performed on a small scale.

- Challenges: Checking these sources can be highly time-consuming, and designers must determine the source's reliability. Furthermore, designers also need to determine if what the participants "say" correlates with their behaviour in reality.

- Possibilities: Operationalizing bigger scale data analysis and discovering user needs quickly.

> Vehicle-related data:

There are multiple data sources within this group, most of them quantitative and complex to categorize. Nevertheless, a classification based on the information obtained in the interviews is presented below.

» Data directly retrieved from the vehicle itself:

Information about the vehicle status (its components and parts), the environmental conditions, the driving behaviour or location. The sensors that collect this data can operate in two different ways:

- Uni-directional signals recorded by the Telematics Control Unit (TCU), like vehicle data health.

- Bi-directional signals, this type of source allows Ford to send and receive inputs from the car; therefore, new software features can be tested.

» **Data from services related to the vehicle:**

The two main sources are Ford Pass and Sync Analytics. The first one is an application that aids the users with tasks related to the vehicle (e.g., checking if the vehicle is locked). The second is the car's infotainment system and collects information about the usage of vehicle features and the media.

» **Data from other departments:**

In this group, we can find information derived from other departments like the Ford Customer Service, the vehicle production or the sales.

The following variables have been taken into account to understand the complexity, challenges, and possibilities of the data available:

- Availability: All the new vehicles have a modem implemented to collect the main data functionalities and store them in a cloud infrastructure. If the designers would like to research older models, external PID (Plug-In Device) modules should be implemented under the approval of the car owner.

- Privacy: For customers' privacy, the company complies with the European regulation. The level of approval by the vehicle owners/users will affect the possibilities of use by the design team. At the moment, the only data that is directly collected and cannot be identified to a vehicle is SYNC

- Context: most of this quantitative data can be considered "thin" and lacks context. For instance, the sensor on the car can tell the designers the door opened, but the question is how to relate this event to others to make sense of it.

  • Challenges: identify which data source will be optimal for the design process.

  • Possibilities: understand large-scale patterns and trends.

In conclusion, the figure on next page summarizes the state of the data use and the different sources available for the Ford team based on the Bornakke and Due (2018) classification (Fig 9). These authors organise data on a double dimension matrix: "volume" and "context". Volume ranges from "small" to "big", depending on the number of instances. The second dimension is measured from "thin" -little context linked to them- to "thick"-high context complexity. Context complexity is a characteristic "which enables the researcher to reflect upon how and why people do what they do" (Bornakke and Due, 2018) and therefore understand users' behaviours and needs.

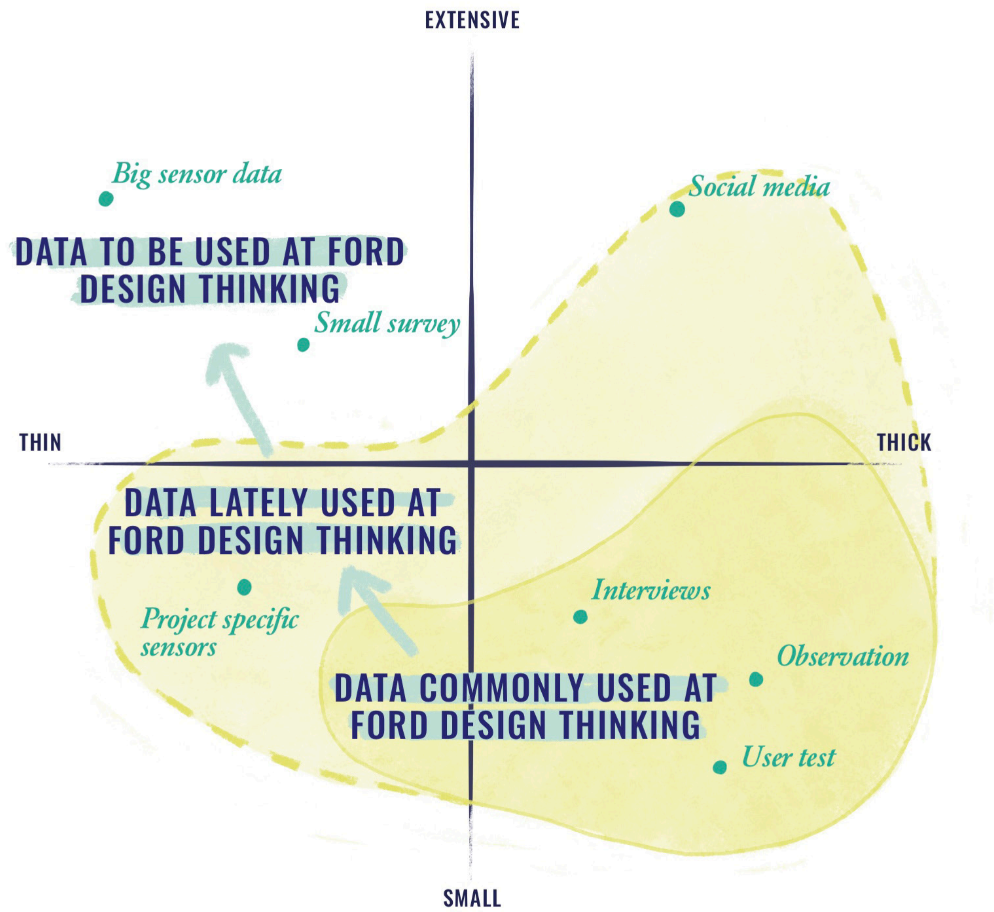


Fig 9. State of data use within Ford's design thinking process using the data classification system from Bornakke and Due (2018). Based on my research and Jansen (2021)

- The Smart Vehicles team has access to multiple data sources, which can be overwhelming specially when referring to data sources the team is not experienced with like big data.
- The implementation of more types of data in the Ford's design thinking process is a great challenge that is taking time .
- The team is most experienced with the data collected by themselves both qualitative (like interviews or observations) or quantitative (from specific sensors installed in the vehicles)
- There main challenges to use vehicle related data with low context complexity is the difficulty to relate it with other data sources due to privacy

2.5. DESIGN PROCESS & DATA USE

As explained before, since 2017, one of Ford's objectives has been to implement a Design Thinking process with a human-centred approach. To reach these goals, the company collaborated with IDEO to create a theoretical framework based on the desired mindset: a working process model (Fig 10) and the methods employees might need.

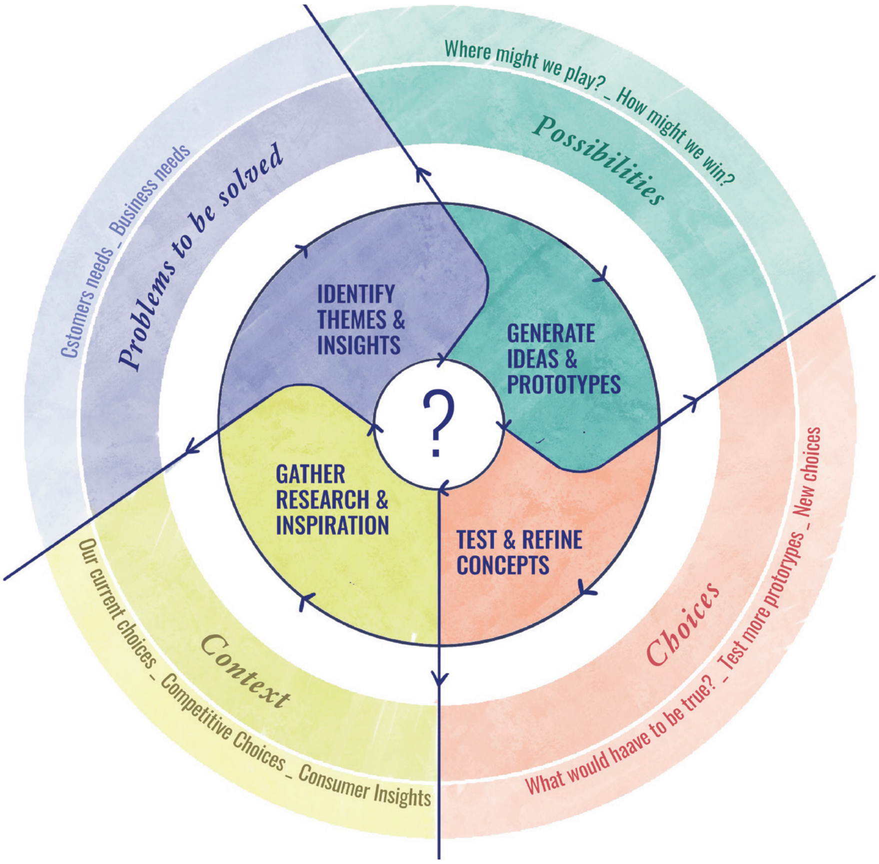


Fig 10. Ford's design thinking process developed in collaboration with IDEO Adapted from Ford.

In practice, implementing these changes needs time, experience, and adaptability to the employees' differences observed in section 2.3. Consequently, the IDEO model serves as a general guideline for the Smart Vehicles Concept to develop their projects which can be adapted to each project's requirements.

> ***"There is not a clear and defined way on how we create ideas, we learn by doing (...) We adapt constantly"*** (Interviewee Smart Vehicles Concept, 2021)

The projects generally start with the definition of the central question. Next, a phase to gather information and inspiration about the context with mainly qualitative research, including interviews, surveys, observations, or social media exploration. Afterwards, in the second phase, the material is shared with the entire team through multiple creative sessions to discover insights and identify the current problems. Finally, based on this knowledge, the team enters another round of sessions to generate ideas and/or prototypes that will be evaluated and refined later. Thus, the process is not linear but cyclic with multiple iterations.

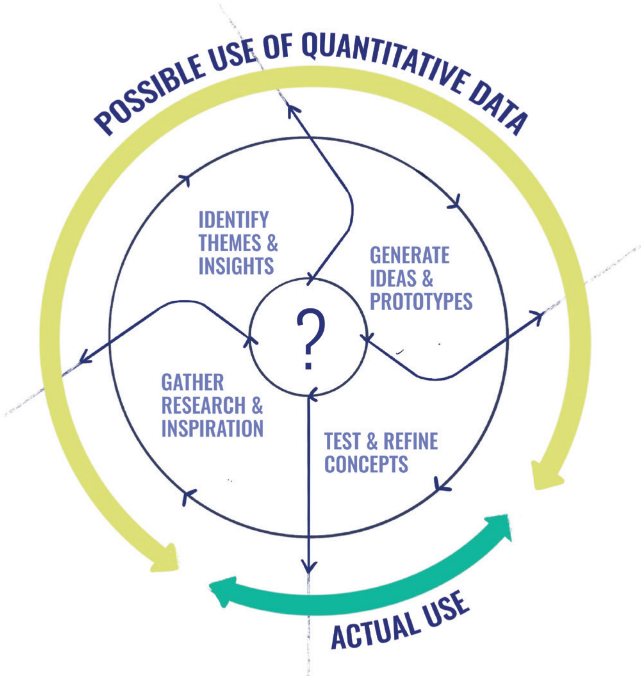


Fig 11. Ford's design thinking could benefit from using the quantitative data in more stages of their creative process.

In this process, the role of data is critical as the first stage (Gather Research) is mainly based on it. The Smart Vehicle Concepts team is experienced in using qualitative data through all the steps. For instance, they can look in an open data source like You-tube videos to understand the user behaviour, recruit participants to interview based on their observations, co-create solutions, and receive feedback on their concepts. By contrast, the use of quantitative data seems to be restricted to a small scope (as seen in Fig 11); and to the later stages for verifying decisions and avoiding design biases. Therefore, there is still a significant potential to use quantitative data, especially considering the variety of data that the design team could access.

- The team learns and grows by doing and gaining more experience over time. In any case, the iterative steps of the Design Thinking process are generally followed as outlined in the figure.
- The design team uses primarily qualitative data (collected by themselves in surveys, interviews, user tests or others); especially in the initial stage of the creative process as inspiration and as feedback during idea/prototype generation.
- The quantitative data (retrieved from the vehicles or a specific PID) is mostly used to analyse a concept's results and as a base for the following projects.
- There is a potential for the development of new ways to use quantitative data.

2.6. QUANTITATIVE DATA PROCESS & VISUALIZATION

As the use of big data is relatively recent in the company, the different involved teams are still defining and adjusting the data approach as efficiently as possible. Nevertheless, by now, the ideal process as determined by the department CDF is as follows:

Fig 12. Current Ford's data process model. Adapted from Ford

- 1. Request Clarification: An employee from any department (like Marketing or Design) contacts CDF. Together they evaluate what data the employee needs and the reasons behind it. To help in this task, the CDF has created a general template to fill up by the requester. This document is not so extensive and asks for as specific as possible descriptions, being the first one and most important: "What results would you like to see presented?
- 2. Feasibility Analysis: the sources, signals, or sensors necessary to obtain the data that answer the questions defined in the previous step are defined. Depending on the request, it could be just one source of data or a combination.
- 3. Data Privacy Approval: the plan defined in the previous step must go through a double approval process by the Governance Board, a committee external from GDIA, CDF and SVC.
- 4. Pilot Data Operations: A pilot study starts to verify that data extraction can be carried out. This study is done on a small sample of vehicles.
- 5. Pilot Data Analysis: With the data obtained, a model is generated that allows its analysis and, if necessary, its visualization that can be scaled to the total population in the future.
- 6. Data Deployment: If everything proceeds correctly, the process to deploy the entire sample begins.
- 7. Data Extraction: The data of the established period are extracted. This period can generally never exceed six months; this time range offers enough time to extrapolate conclusions without accumulating costs.

- 8. Request Report Out: The report is generated based on the approved Feasibility Analysis and the Pilot carried out.

- 9. Request Completion: The employee who requested the data obtains his report, which could be from an Excel file to a complete visualization panel.

As said before, this is a general model applied inside Ford, which is still evolving. Hence, it could be an excellent opportunity to adapt the data process to the different needs of each department.

The design team faces multiple challenges, and I consider the most important the following ones:

- The data departments, CDF and GDIA, ask for the specific results expected or a determined hypothesis to verify (e.g. the data request template). This mindset, in general, collides with the designer's attitude, which, as Jansen (2021) identified, usually does not have one straightforward question yet because they are still determining the problem.

  "So basically, first of all, we need to have the right question" (Interviewee GDIA, 2021)

  "We do not always have a question we want to use the data with, because for example, we want to discover patterns, and how do you ask a question to discover a pattern?" (Interviewee designer SVC, 2021)

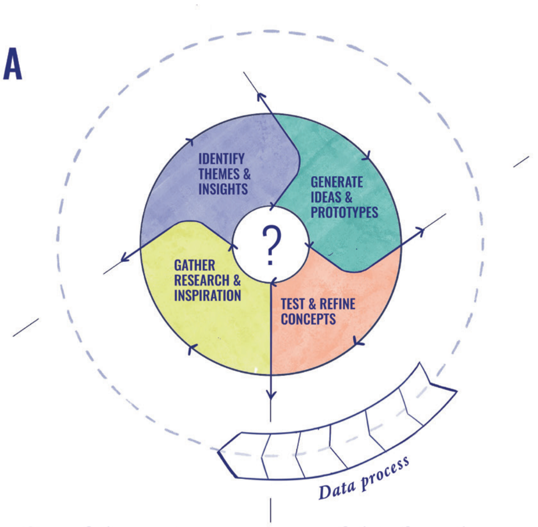
- There is no standard guideline on which software should be used at any stage of the process. It depends on both the employee and the source of the data. The GDIA uses mostly Spark or Alteryx to clean, evaluate, and process raw data. The SVC team relies in Excel or MATLAB. The use of multiple softwares can difficult sharing information across departments, spending more time converting the different files or having compatibility issues among others.

- Concerning data privacy:

  - In general terms, the aim for the data use should be established from the beginning; furthermore, only the questions approved by the Governance Board can be answered. "If you want to answer new questions, you have to go through the whole process again" (Interviewee CDF, 2021). This clear statement hinders an open exploration of the dataset from the design team.

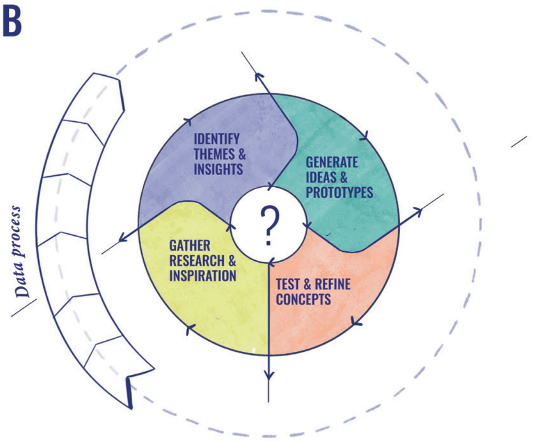
  - Only the specified and approved results can be shared between the also determined audience. Employees who analyse the data cannot share other details that they may have observed in the database.

A



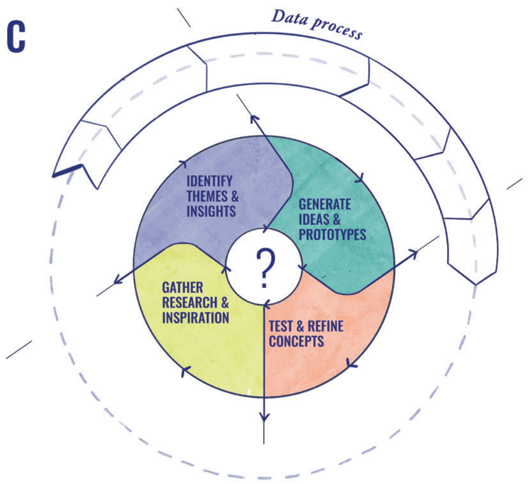
One of the most common uses of data from the vehicles, is to test the design concepts and afterwards use that input as the base for next projects.

B



The data process can be started at any phase of the creative process. In this case, the data will be used as creative material to Identify themes & insights.

C



Designers have to consider the time required for each specific data request. As seen in the image, the designers can get the results after a long period.

- The data cannot be identified with a specific user. Therefore, the design team cannot track back and obtain more information about the context or develop a particular user test. Understanding how to preserve privacy rights and bring the most out of the data is an excellent question for designers and data scientist

» The data specialists are not involved in the design team's ideas behind the requests, especially after the results are delivered. However, the CDF has already noted the necessity to extend their participation. In addition, GDIA also indicated their ambition to improve the communication with the design team (as reported in the section ….)

"I do not know whether the result suits his expectation or if it is something completely different. We are not concentrating on what the requester is really doing afterwards with the data that we provide, and we are realizing we need to control this to improve" (Interviewee CDF, 2021)

» Involving quantitative data results in a long process requiring many resources and steps that can discourage the design team from submitting a data proposal.

"I just want to say that it would be possible to create more data, but they're not coming from nowhere. That's a huge amount of effort to collect data." (Project manager interviewee from SVC, 2021)

These challenges are present when the design team wants to use data related to the vehicles, no matter the creative process stage. In general terms, the relationship between the creative process and the design process can be understood as in the left figure (Fig 13). In many cases, the time of the data process can be longer than the time estimated for the different stages of the creative process. As a result, the SVC department has to adapt their project timing or decide not to use data that requires GDIA intervention.

Fig 13. Different possible scenarios when combining Ford's design thinking and data process.

2.6.1. THE USE OF DATA VISUALISATION

As said in section 1.3, this project aims to discover the possibilities of data visualisation to improve the use of data as creative material. Hence, understanding how Ford is currently using data visualisation is a crucial step for developing new opportunities.

Data visualization has a minor role in the data process, if used at the end of it (check Fig 12), and there is no established department responsible for its performance. If used, it is intended to present a determined result, to communicate something already known to other employees, as I observed in the interviews:

> ***"When you create a visual you already should have an idea about what you want to show and how you want to show it"*** (Designer interviewee from SVC, 2021)

In general, to convey the results, simple static visualizations are used like histograms or pie charts.

> ***The standard dashboards are just very straightforward bar charts about, for example, the distribution of a trip length.*** (Designer interviewee from SVC, 2021)

In terms of software use, it mostly depends on the employees' preferences:

- QlikView is the software standardized by Ford to create official publications, presentations, etc. Therefore, there are style guideline templates established by a team in Ford North America. Furthermore, it allows the creation of dashboards that visually display all the results. The main problem is that to be licensed for this program, any employee at Ford must pass a stipulated course and obtain accreditation which is not common, and mainly just GDIA has access to it.

- MATLAB and Excel. Most of the time, when the design team just requests the data as a table, they employ one of this statistical-mathematical software. Unfortunately, both programs are not explicitly created to perform data visualization and probably limit the opportunities that data visualization offers to analyse the data.

- Tableau. Some employees in the design department have noticed the need to use higher-level visualization software than Excel or MATLAB. Due to the hazards of getting a QlikView accreditation, they have started using Tableau. However, it requires a personal effort as designers are generally not trained and prepared to work with big data and its visualization (Davenport et al., 2019).

In conclusion, as far as I have observed, data visualization, in general, is not considered a tool to discover and analyse datasets.

- Ford has a defined process for the use of big data, which is in evolution, and can be modified and adjusted to obtain greater efficiency.
- Data privacy, communication, collaboration, and time are big challenges for using big data in Ford's creative process.
- Currently, data visualization generally plays a minor role in Ford's data process and is not considered a tool for discovering and analyzing data sets.

3 THEORETICAL EXPLORATION

Now that we have better understood the challenges and opportunities currently present in the data-enabled design process at Ford, the next pages aim to introduce the most relevant research findings of creativity and data visualization on which this thesis is based. By reading the following explanations you will be ready to dive into the next chapters and consequently have a better understanding of the decisions taken along the process.

3.1. Research approach _ *Page 34*

3.2. Creativity _ *Page 36*

3.3. Data Visualization _ *Page 45*

3.4. The intersection: Crossing Bridges _ *Page 50*

3.1. RESEARCH APPROACH

Based on the initial research question of this project: How can the Ford Smart Vehicles Concept team use Data Visualization as an exploration tool to facilitate their creative process, leading to innovative insights? The following sub-question led this research stage:

How can data visualization support the creative process of designers?

The approach included different usual activities of a research process, from theoretical ones like the search and review of state of the art to practical ones like practising with data visualization tools like Tableau and Voyager 2[1].

In the first phase, I explored the topics of creativity (section 3.2) and data visualization (section 3.3), which led me to conclude with eight possible opportunities that data visualization offers for the creative process, which I named Crossing Bridges and are collected in section 3.4. After selecting Exploratory Inquiring[2] as the most interesting Crossing Bridge to develop for Ford (explained in chapter 5), I carried out a second phase of research to understand the theoretical roots of this design direction through two fields Exploratory Data Analysis (3.3.4) and Questions in Design (3.2.5) .

Instead of following a chronological order, this chapter gathers the different topics researched into two significant clusters: Creativity and Data Visualization; and concludes with my contribution: the Crossing Bridges. On the next page, Fig 14 offers an overview of the whole process and identifies the sections and chapters where each topic is tackled.

1 Voyager 2 is a data exploration tool that blends manual and automated chart specification, to support both breadth-oriented exploration and depth-oriented question answering. (for more information check https://vega.github.io/voyager/)

2 In a few words: Using data visualization to facilitate the generation of questions to discover insights and improve the creative process.

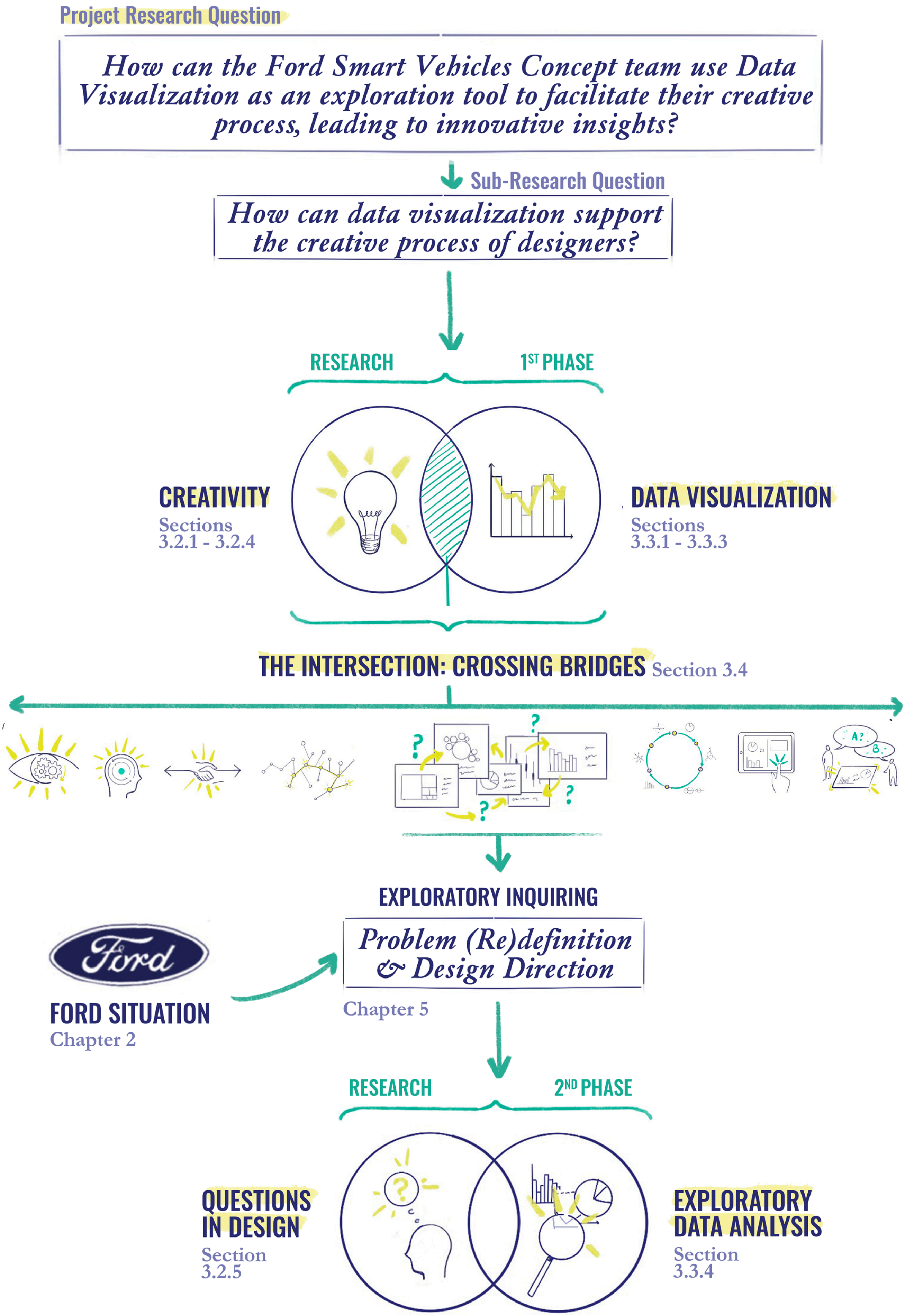


Fig 14. Overview of the research conducted in this project

3.2. CREATIVITY

3.2.1. DEFINITION OF CREATIVITY

Creativity has always intrigued humans; it is something almost all of us desire to achieve, even though it is tough for us to define what it exactly means. Scholars have widely diverged on its definition, in part because its meaning depends on the viewpoint chosen to explain it. In 1961, Rhodes distinguished four main perspectives to tackle creativity. This classification, the 4P's model (), is currently one of the most relevant and referenced:

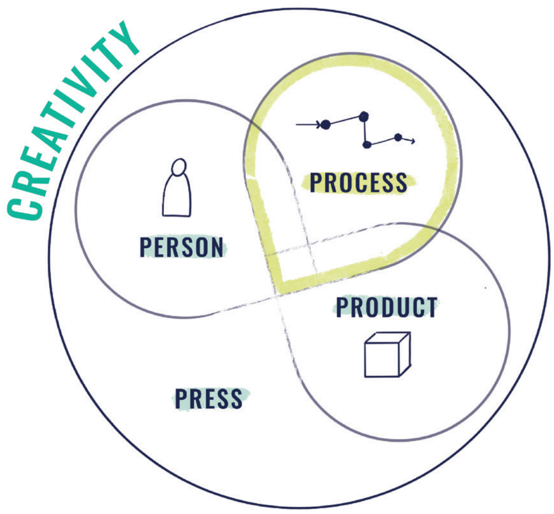


Fig 15. The 4P's model interpretation from Rodes (1961)

» Person: covers "information about personality, intellect, temperament, physique, traits, habits, attitudes, self-concept, value systems, defence mechanisms, and behaviour."

» Process: applies to "motivation, perception, learning, thinking, and communicating".

» Press: refers to "the relationship of human beings and their environment" (Rhodes, 1961)

» Product: creativity amounts to the outcome: the product or idea.

Taking into account that the interest of this project is on improving Ford's creative process, we are going to focus on the "creativity" process perspective and use one of the definitions most referenced:

"Creativity is the process that leads to novel and useful solutions to given problems" (Amabile, 1996)

3.2.2. THE CREATIVE PROCESS

Early researchers mainly explained the creative process through linear models, which well-known creativity experts like Torrance (1988) or Plsek (2006) claim are based on Wallas' one.

The model attributed to Wallas (1926) (Fig 16) explains creativity concerning problem-solving in four stages:

- Preparation: analysis of the problem and its context; frequently constructing a problem statement.
- Incubation: a period where the problem solver is not consciously thinking or working on it.
- Insight / Illumination: a solution arises, typically resulting from unexpected associations.
- Verification: the solution is critically evaluated and elaborated.

Although these phases are continuously referred in subsequent literature, the creative process is no longer understood as a consecutive chain of action but rather as a cyclic and iterative process (Tardif and Sternberg, 1988). This would mean that the whole process results from small creative thinking loops that add to each other (e.g., Amabile, 1983; Cross, 1997; Crilly, 2010).

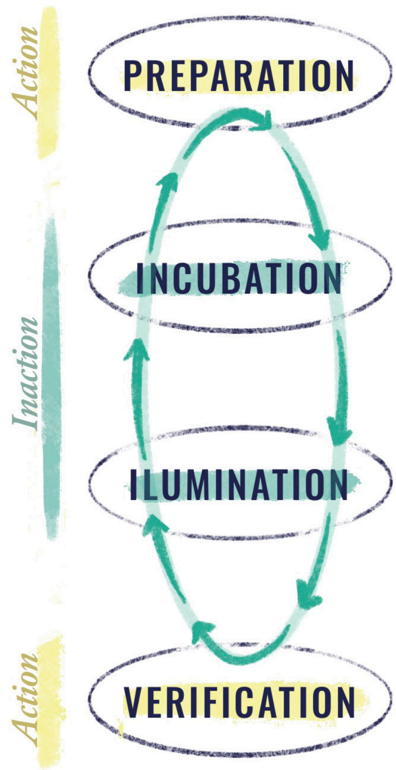


Fig 16. The creative process according to Wallas (1926)

> Creativity can be understood as a process to generate useful and novel solutions to a specific problem.

> The creative process is iterative and cyclic, involving both unconscious and active stages.

3.2.3. CREATIVITY, A CONSTANT PLAY BETWEEN SOLUTION AND PROBLEM

As a whole, the creative process could be understood as developing and refining iteratively together both the formulation of the problem and the solution space (Dorst, 2019) (Fig 17).

This notion, referred to as co-evolution, helps to explain also the creative practice of design. Designers do not first understand the problem and then develop a satisfactory solution. Instead, designers modify their problem statement as their exploration of possible solutions helps them better understand it until they find "an idea" that adequately connects the problem and solution (Dorst & Cross, 2001).

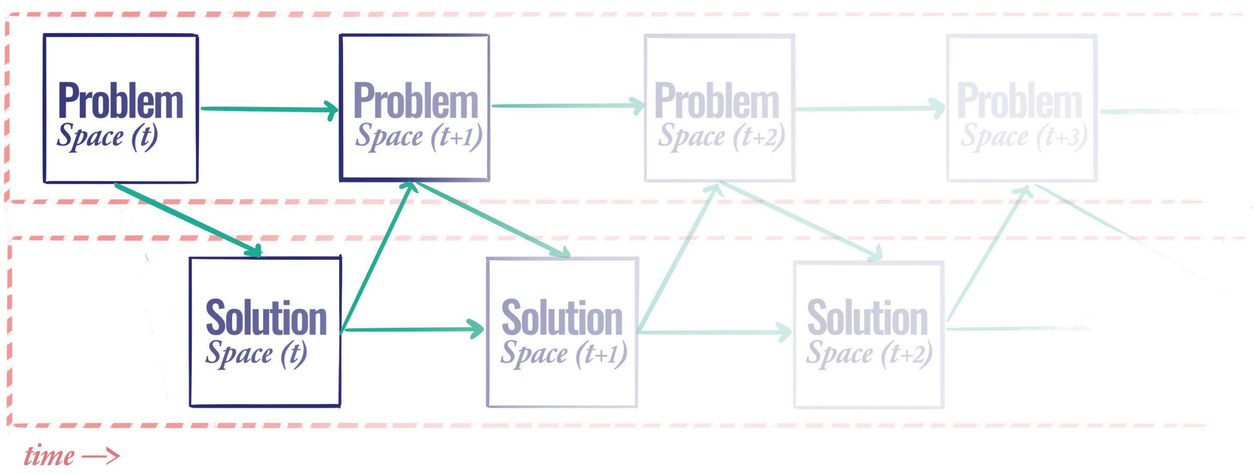


Fig 17. The Co-evolution model adapted from Dorst & Cross (2001)

3.2.4. THE IMPORTANCE OF KNOWLEDGE IN THE DESIGNERS' CREATIVE PROCESS

The co-evolution model states that there is an "interchange of information between both spaces" (Dorst, 2011). This "information" flow is constant throughout the whole creative process and is critical for both solution ideation and problem formulation.

The effects of knowledge on the interplay between problem definition and solution generation have led to Hui et al. (2020) to revise the Dorst's model and create the Triple Helix Structured Model. In this model, the solution and problem are complemented with the knowledge space. As seen in the figure, the constant interaction (or "mapping") between the three spaces can be understood as a spiralized co-evolution that feeds the design process.

Knowledge is pivotal for the majority of the creative processes models (Lubart, 2001) and the base for this knowledge is data (Ackoff, 1989).

In this project, we will consider as a fundamental premise the critical role of knowledge to shape the solution and problem space in a creative process.

- The creative process is a constant interplay between problem framing and solution exploration!
- Designers need the knowledge to develop and refine the solution and problem space.
- Therefore, specific data can be the fuel of the knowledge space giving the designers insights about the user and its context.

3.2.5. QUESTIONS IN DESIGN

One of the key limitations of creativity is asking the wrong question.
Claudia Kotchka, Innovation and Strategy Advisor and IDEO Fellow

As explained previously, creativity involves knowledge, and questioning is one of the most potent reasoning mechanisms humans have to discover insights. As Berger (2014) indicates, "questioning enables us to organize our thinking about what we don't know".

Questions have long been considered powerful tools to help increase creativity in different fields such as education or psychology. Still, the work of designers also relies on the questions they formulate (Eris, 2004), and the interest in the role of questions in the design process is relatively recent.

The first wave of research was motivated by the need of developing digital tools for designers to retrieve information (Kuffner and Ullman, 1991; Gruber and Russel, 1992; Baya,1996). These pioneering researchers suggested different question taxonomies, identifying 4, 14, and 11 types of question categories, respectively, based on the request for information. In general, these classifications have not been particularly relevant for further research as they are very simple and incomplete (Aurisicchio et al. 2006). Nonetheless, Baya's observations on designers' behaviour paved the way for investigating question asking as an inquiry process. This author exposed that designers not only do not have a predefined set of questions but also do not form questions randomly; designers form new questions after reflecting upon information received as an answer to their previous questions (Baya, 1996)

After this technological-driven approach to questions, subsequent studies switched the focus on understanding the inquiry process on the design activity (Eris, 2003; Dym et al. 2005; Eris, Sheppard & Kwan; 2007; Ahmed & Aurissicchio, 2007; Grebici et al., 2009; Aurisicchio et al. 2010, Cardoso et al. 2014, Cardoso et al. 2020, Hurst et al. 2021). The most relevant contributions for this thesis are mentioned below, which is mainly based on the work of Eris.

Eris (2003) advocates that design is inquiry-driven and elaborated a complete design questions classification up to date, extending on relevant taxonomies from various disciplines: philosophy (Aristotle), education (Dillon, 1984), artificial intelligence (Lehnert, 1978), cognitive psychology (Graesser, 1994), and design research (Kuffner 1990, Baya 1992) (see Table 2 on the next page).

In his framework, the design process is an interplay between low-level and high-level questions. The first group (LLQ) comprises Factual Questions, which require less cognitive effort and aim to request information related to the attributes or the existence of a subject, object or phenomenon. For example, imagine a designer asking: *What is the material of this product?*; his question indicates the aim to retrieve a feature of a specific component. Despite

ARISTOTLE	DILLON	LEHNERT	GRAESSER	ERIS	
Existence (Affirmation)	Existence	Verification	Verification	Verification	Low Level Questions (LLQ)
	Instance				
Nature (Essence)	Substance		Definition	Definition	
			Example	Example	
Fact (Attribute/ Description)	Character/ Description	Feature Specification	Feature Specification	Feature Specification	
		Concept Completion	Concept Completion	Concept Completion	
		Quantification	Quantification	Quantification	
	Function	Goal Orientation	Goal Orientation	Rationale/ Function	
	Rationale				
	Concomitance	Disjunctive	Disjunctive	Disjunctive	
	Equivalence		Comparison	Comparison	
	Difference				
		Judgamental	Judgamental	Judgamental	
Reason (Cause/ Explanation)	Relation		Interpretation	Interpretation	Deep Reasoning Questions (DRQ)
	Correlation				
	Conditionality & Causality	Causal Antecedent	Causal Antecedent	Causal Antecedent	
		Causal Consequence	Causal Consequence	Causal Consequence	
		Expectational	Expectational	Expectational	
		Procedural	Procedural		
		Enablement	Enablement	Enablement	
				Enablement	Generative Design Questions (GDQ)
				Method Generation	
				Proposal/Negotiation	
				Scenario Creation	
				Ideation	

Table 2. Comparison of the different taxonomies reviewed by Eris (2004)

their name, these questions do not have a "low" value, but they precede the high-level questions.

There are two types of questions on the higher level: Deep Reasoning Questions (DRQ) and Generative Design Questions (GDQ). The first ones reflect convergent thinking and are used to understand facts. In particular, questions that belong to this group could be: *Why this product works like that? What are the effects of this product feature on the users' behaviour?* The second ones, GDQ, are natural from divergent thinking processes and aim to create possibilities from facts. For instance, when refining a product, a design team could ask themselves: *How can we keep it from behaving in this unexpected way?* as a way to generate different solutions to prevent an undesired result. Eris was the first author to identify these questions, typical in design tasks and directed to open up the solution space as Generative Design Questions (GDQ).

The next Table 3 offers and overview of the main characteristics of each group of questions (for a description and examples of each subcategory check Appendix: Eris Taxonomy completed with Personal Examples)

	LOW LEVEL	HIGH LEVEL	
	FACTUAL QUESTIONS	DEEP REASONING QUESTION	GENERATIVE DESIGN QUESTION
Purpose	Retrieve missing info or confirm information	Understand and explain facts	Create possible opportunities
Cognitive Mechanism		Convergent Thinking	Divergent Thinking
Answer	The answer is known, if not by the subject, by someone else	The answer is known, if not by the subject, by someone else	For any given question, there exist, multiple alternative known answers as well as multiple unknown possible answers that are yet to be created

Table 3. Summary of the question taxonomy proposed by Eris (2003)

In his studies, the author observed a high correlation between the use of DRQ and GDQ and the design team performance. This observation leads him to suggest that each group of questions serves different purposes, and its use depends on the design phase. In his experiments, he observed a higher frequency of particular subtypes of questions depending on if, at that moment, the designers were conceptualizing, implementing or assessing (Table 4). In this project, we can consider the Conceptualization stage as it "involves users' need finding, requirements definition, and idea generation" (Eris, 2004).

Lastly, Eris indicated that the different communication mediums (e.g., sketches, digital interfaces or writing documentation) used by designers would create particular questions-posing opportunities. In particular, he proved how the access to "existing artifacts and prototyping hardware" (in his experiments:

	DESIGN PHASE		
	CONCEPTUALIZATION	IMPLEMENTATION	ASSESSMENT
Low Level Questions (LLQ)	Verification	Verification	Verification
	Disjunctive	Disjunctive	Disjunctive
	Concept Completion	Concept Completion	Concept Completion
	Feature Specification	Feature Specification	Feature Specification
	Quantification	Quantification	Quantification
	Definition	Definition	Definition
	~~Example~~	~~Example~~	~~Example~~
	Comparison	Comparison	Comparison
	Judgmental	Judgmental	Judgmental
Deep Reasoning Questions (DRQ)	Interpretation	Interpretation	Interpretation
	Procedural	Procedural	Procedural
	Causal Antecedent	Causal Antecedent	Causal Antecedent
	Causal Consequence	Causal Consequence	Causal Consequence
	Rationale/Function	Rationale/Function	Rationale/Function
	~~Expectational~~	~~Expectational~~	~~Expectational~~
	~~Enablement~~	~~Enablement~~	~~Enablement~~
Generative Design Questions (GDQ)	Enablement	Enablement	Enablement
	Method Generation	Method Generation	Method Generation
	Proposal/Negotiation	Proposal/Negotiation	Proposal/Negotiation
	Scenario Creation	Scenario Creation	Scenario Creation
	Ideation	Ideation	Ideation

Table 4. In this general matrix, the coloured categories represent the questions Eris observed a relative distribution. The non-coloured were present; however, their low incident rate led Eris not to consider them relevant for that design phase. The crossed categories were types of questions that Eris could not observe in any experiment.

constructing bloks) modifies the questioning process, increasing the number of DRQ and helping to regulate the divergent thinking process. These findings may suggest that data visualization can enable a specific type of questioning process.

Posterior studies by the same author confirmed that question asking changes with experience: design students increase the use of generative design questions as they progress in their undergraduate education. Project-based learning methods were considered the possible reason for this development (Eris et al., 2007).

In the field of design education, Dym et al. (2005) demonstrated the importance of teaching effective queries for the design process, precisely questions that promote divergent thinking.

Currently, the research carried out on practical cases is mainly focused on identifying the questions with the information required by the designer. Furthermore, these studies also concentrate on product design (Ahmed and Aurisicchio, 2007; Aurisicchio, Bracewell & Wallace, 2010). These circumstances limit the application of its conclusions to other still unexplored practical cases, such as in services' design or projects that could need a more broad exploration.

The latest studies on the field explore the role of design questions in design reviews (Cardoso et al. 2014) and peer feedback sessions (Cardoso et al. 2020, Hurst et al. 2021). They identify the questions needed for a more compelling experience based on Eris taxonomy. These examples reaffirm the benefits of understanding the design questioning process at all stages of the creative process.

- Design is highly based on question asking, there are different types of questions and a good balance of questions, which promote convergent and divergent thinking, improve design teams' performance.
- There are two main types of questions: low questions, which require less cognitive levels and high level questions, which happens by convergent or divergent thinking.
- The use of high level questions (GDQ and DRQ) can be improved by the presence of "hardware"
- Is relevant for designers to learn to formulate high level questions and studies have shown that it is possible.
- Currently, the role of questioning in design has not been sufficiently explored.

3.3. DATA VISUALIZATION

3.3.1. DEFINITION OF DATA VISUALIZATION

Data Visualization, also sometimes referred to as Information Visualization (InfoVis) or Scientific Visualization (Ziemkiewicz & Kosara, 2007), can be understood as "the use of computer-supported, interactive visual representations of data to amplify cognition" (Card, Mackinlay, and Shneiderman, 1999). As cognition refers to "the mental process of acquiring knowledge and understanding through thought, experience, and the senses" (Oxford University Press, n.d.), data visualizations are graphic aids to transform data into knowledge. InfoVis systems are the perfect complement to humans' capabilities but not wholly replacing humans. Machines are still far from approaching human abilities in synthesizing new knowledge, hypothesis formation, and creative modelling.

The process of drawing insights and conducting investigative analysis is still clearly in the realm of tasks best left to humans. (Munzner, 2014). In order to obtain these insights, data visualization supports our cognition by exploiting the human visual system, "our natural means of perception" (Bertin, 1983). In this respect, around 50-80 per cent of brain resources are dedicated to our vision (Krum, 2013), and our brain tends to process visual information before any other sense (Pike et al., 2012).

Hence, Data visualization is a complex activity since it involves three different functional spaces: a computational space, an interaction space and a mental space. In addition, to validate and improve visualization design, researchers need to understand and define how users interact with visualization systems under realistic scenarios.

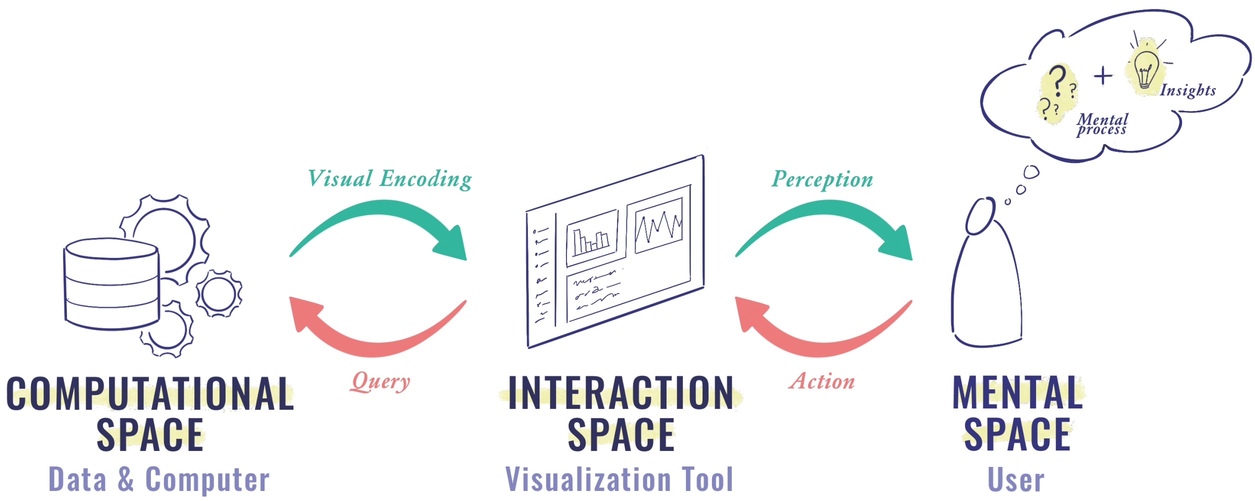


Fig 18. A data visualization system involves three functional spaces: a mental space, an interaction space, and a computational space adapted from Sedig et al. (2012)

3.3.2. THE COGNITIVE MECHANISMS TO VISUALLY EXTRACT INSIGHTS

In general, the processes by which information visualization users seek and obtain knowledge are not well understood (North, 2006; Yi et al., 2008; Reda, 2016). In this project, the model developed by van Wijk (2005), refined by Keim et al. (2008) and later improved by Green et al. (2009), will be used as it offers a simple overview (Fig 19).

As seen in the figure, in this model, the starting point is data (D), which is transferred to a visualization system (V) and perceived as an image (I) by the user. Next, the perception system (P) extracts knowledge (K) over time (t) which ideally leads to further exploration and analysis (E) through interactive changes to the visualization specification (S).

This model shows that obtaining knowledge depends both on the users' perception of the image and the different opportunities that the user can use to explore the visualization (for instance: select outliers, change the colour coding or zoom for details)

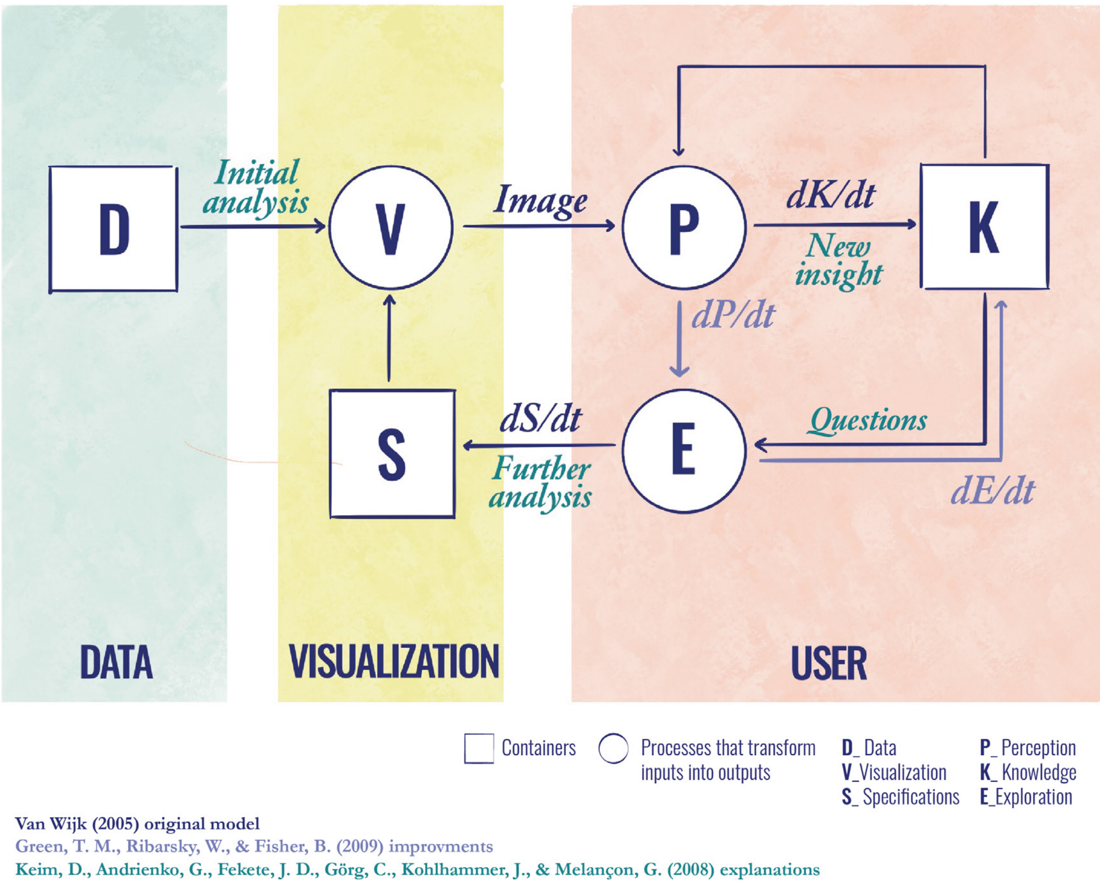


Fig 19. The sense making loop of visual analytic adapted from Van Wijk (2005) incorporating the improvements from Keim et al. (2008) and Green et al. (2009)

3.3.3. THE COMPLEXITY OF REPRESENTATION IN DATA VISUALIZATION

An effective data visualization will assist users on the cognitive loop previously explained. However, creating a Data visualization is not a simple task. Firstly, the type of visualization is determined by the data types (quantitative, qualitative,...). Furthermore, the main dilemma is that a particular visualization might be accurate for an intended action but not cover other tasks (Munzner, 2014). The visualization creator faces the challenge of choosing an appropriate visual encoding that will also allow the user to perform different actions. In this project, I decided to apply the framework developed by Tamara Munzner (2014). Since it helps non-experienced creators understand what data users require to be visualized, why users need to execute specific tasks, and how the visual representations can be composed and manipulated.

The purpose of visualization is insight, not pictures
Card, Mackinlay, and Shneiderman (1999).

- To understand Data Visualizations, we should focus on three different levels: computational, interaction and mental.
- Data Visualizations rely on the human visual sense to transform data into knowledge.
- Data Visualization is a supportive tool, a complement to human cognitive skills!
- The visual representations and interaction possibilities inherently influence how the users obtain insights from the data.
- Creating a compelling Data Visualization is difficult because representations have to be carefully chosen and created to aid a particular combination of tasks and data combinations.

3.3.4. DATA VISUALIZATION AND EXPLORATORY DATA ANALYSIS

One of the most significant values of Data Visualization to understand a data set is performing Exploratory Data Analysis (EDA).

Although data scientists were already using specific techniques in this field, the term EDA was not defined until 1977, when John Tukey published the book Exploratory Data Analysis (Tukey, 1977), as a complementary perspective to the well-extended Confirmatory Data Analysis or CDA ("a statistical analysis designed to address one or more specific research questions, generally with the objective of confirming preconceived hypotheses" APA (n.d.)). EDA is an "attitude, a state of flexibility", a "detective work designed to reveal the structure or patterns in the data" (Tukey, 1980). Applying EDA means answering the question "What's going on here?" when confronted with a dataset (Beherens, 1997), besides focusing on the exploration of the data set by generating new research questions, suggesting empirical relationships on the data, identifying outliers or indicating statistical assumptions between others (Behrens, 1997).

EDA is a philosophy that demands an open attitude for data research, which does not indicate which methods or techniques to use. However, compared to tools such as non-graphical counting or numerical statistical methods, data visualization is revealed as the primary tool (Jebb et al., 2017). The reason is that graphics allow the analyst an extraordinary capacity to reveal patterns and extract information. Moreover, visual examination enables users to quickly discover insights (Tukey, 1980), usually unexpected, complex, deep or relevant information (North, 2006), making it simple to generate questions that are easier to answer and advance faster in the analysis. So, data visualization systems facilitate the flexible and iterative attitude proposed in EDA.

Most data scientists advocate for the combined use of EDA and CDA, as EDA can articulate research questions that, when mature, can be tested with CDA (check). However, this situation makes it challenging to isolate EDA from CDA. Unfortunately, there is a lack of literature published in EDA and formal guides to perform this type of analysis (Jebb et al., 2017).

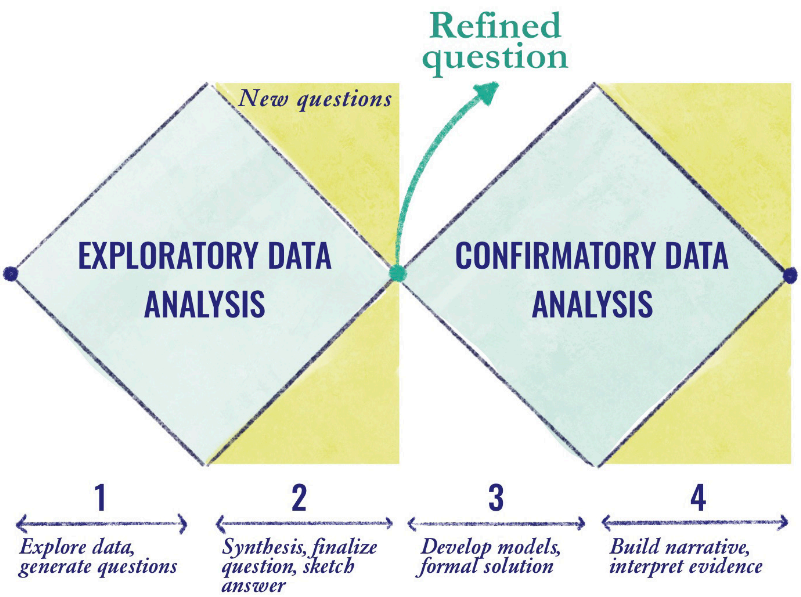


Fig 20. Model of how EDA and CDA can interplay together, inspired by the design thinking "double diamond", an adaptation from Leek, J. & Peng, R. (2020).

As we see, how to generate questions, what methods to use, or which way to go is a path not sufficiently explored yet. So that, for the purpose to contribute to promoting the creation of questions, I reviewed data visualization systems focused on this issue. I noticed that most of the practical approaches are Natural Languages Interfaces (NLI), with flourishing researches in recent years (e.g. Cox et al. (2001); Articulate by Sun et al. (2010); DataTone by Gao et al. (2013); Eviza by Setlur et al. (2016); Articulate 2 by Aurisano et al. (2016)). They aim to support the users by translating their native language hypotheses to a computational query that subsequently will provide a visualization with the requested information. They face multiple challenges like parsing natural language, considering input modalities (e.g., only language-based, language-based + touch, touch + gaze + speech), providing input affordances (e.g., informing people what they can say/ask), explaining system results (e.g., providing feedback on why a chart was shown), among others (Srinivasan & Stasko, 2017).

My conclusion is that the research behind these tools focuses on improving the usability of the visualization systems, trying to generate the desired visualizations that answer the user's questions, but without concentrating yet on supporting the user to develop their questions. Nevertheless, some scholars have already pointed the benefits for users to externalize their hypothesis and the need to provide visualization tools that allow this task (Gotz et al. 2006; Shrinivasan & van Wijk, 2008; Stasko et al., 2008), since this will encourage users in their insights discovery process (Gotz & Zhou, 2009; Ragan et al., 2015).

- EDA is an attitude, a willingness to understand the data by continually questioning its structure.
- EDA focuses on generating the questions rather than finding the correct answer Data Visualization is the primary technique used in EDA.
- Visualization promotes questions that viewers neither initially thought nor knew they could ask, making answering the questions more accessible and faster, moving the analysis forward.
- Defining a good question that drives insights discovery is equal to identifying relationships between particularly interesting or unexpected variables.
- There is no formal guidance to generate questions in EDA or with Data Visualization.
- Externalizing and documenting the questions can help the users to perform a better analysis.

3.4. THE INTERSECTION: CROSSING BRIDGES

Creativity has always intrigued humans; it is something almost all of us de-
The correlations between creativity and data visualization seem self-evident to anyone who delved into both fields; however, there is hardly any literature that explicitly inter-relates these two fields (van Breemen, 2019). Concerning practical applications, only the work developed by Dove and Jones (2013, 2014) was found.

This PhD project consisted of different case studies that tested if data visualization was a potential tool for designers when facilitating a creativity workshop. The results pointed out that data visualization supported the collaboration between participants and enriched their existing knowledge with new insights for the development process of creative ideas.

Regardless of the lack of combined literature, a review of the published literature in each of these fields led me to identify eight possible ways in which data visualization has the potential to support the creative process, which I have termed “crossing bridges”(Fig 21).

Although these crossing bridges overlap each other and are not independent, I believe that exploring each of these areas would allow progress in this field.

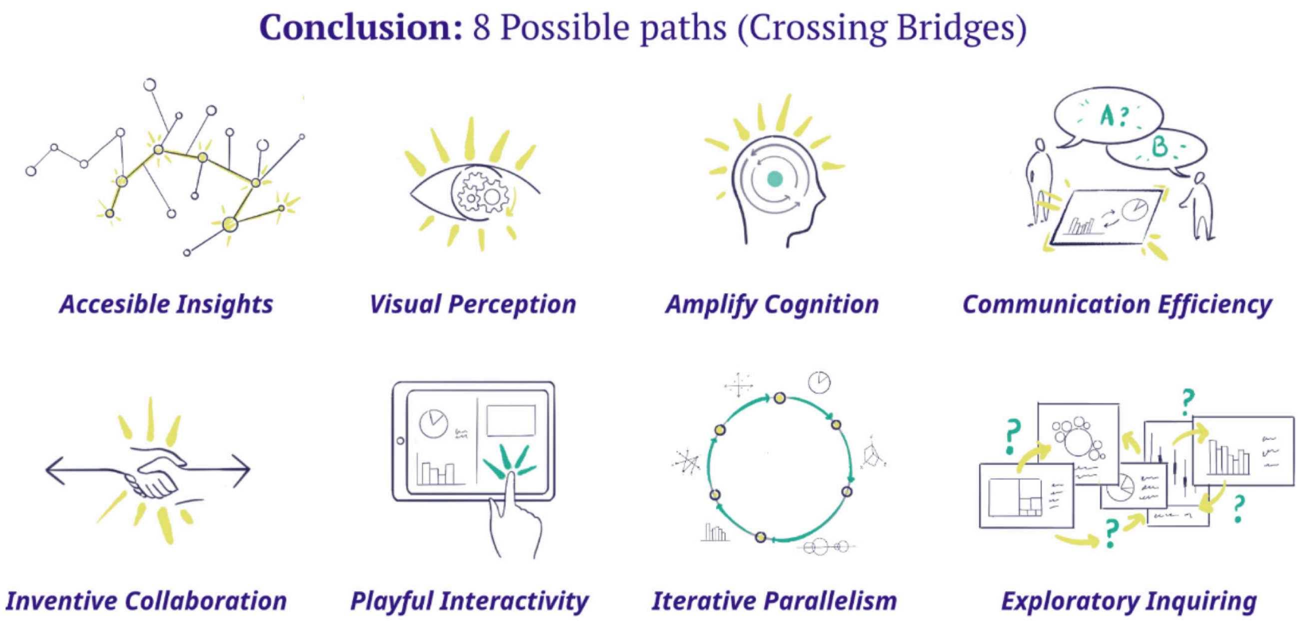


Fig 21. The eight possible design directions obtained through the literature analysis

3.4.1. ACCESSIBLE AND IMMEDIATE INSIGHTS

"Insight is the key of data visualization" (Card et al., 1999), and insights are also pivotal for the majority of the creative process models (Lubart, 2001). In the case of designers, insights contribute to a better understanding of the problem framing.

Furthermore, experts' most underlined aspects are that data visualization offers the opportunity to accessibly gain insights into large datasets (Ware, 2005). Therefore, using data visualization not only could significantly increase the speed and quality of the insights achieved (Batch and Elmqvist, 2017) but can give a general dataset overview and facilitate the decision of the research focus (Jewell et al., 2014)

3.4.2. VISUAL PERCEPTION

From all the senses used to perceive external inputs, vision is the most powerful of all. Humans are used to communicating stories visually because it is the most direct way to carry information to the brain.

Even if almost all individuals have a visual tendency, designers have been highly recognized as primary visualizers (Mednick, 1962; Malaga, 2000). Searching for pictorial stimuli is almost an unconscious inclination for designers, and therefore, is an essential preference during idea generation (Hanington, 2003; Henderson, 1999; Muller, 1989).

3.4.3. AMPLIFY COGNITION

Visualization has been identified as one of the principal cognitive operations employed in the creative process, besides: Combination, Remote Association, Retrieval, Analogy mapping and Abstraction (Ware, 2019)

Concurrently, data visualization has been proved to support users' cognition (e.g., Gershon et al., 1998; Keller and Tergan, 2005; Fekete et al., 2018). Visuals serve to reduce cognitive load, especially for complex task requirements like the case of creativity. Multiple researchers have focused on discovering how external visuals assist cognition; one of the most referenced is the study by Card, Mackinlay and Shneiderman (1999). These authors identified multiple cognitive activities supported by Visualization, prioritizing the following ones:

» Increasing memory and processing resources available
» Reducing search for information
» Enhancing the recognition of patterns
» Enabling perceptual inference operations
» Using perceptual attention mechanisms for monitoring
» Encoding information in a manipulable medium

3.4.4. EXPLORATORY INQUIRING

Creativity involves knowledge, and one of the most potent tools the human has is questioning. Inquiries allow the thinker to discover possible solutions. Likewise, design can be question-based, as Eris (2004) affirms. However, the quality of these questions influences the novelty of the insights and, therefore, the solution creativity level.

In this inquiry process, data visualization can be an invaluable tool. In the words of Fekete et al. (2012), "InfoVis systems, on the other hand, appear to be most useful when a person simply does not know what questions to ask about the data or when the person wants to ask better, more meaningful questions".

3.4.5. COMMUNICATION EFFICIENCY

Creation and design have a wide range of new opportunities if they take advantage of nowadays data revolution, significantly on insights about user habits, desires, and interactions (Marti, Megens and Hummels, 2016). Therefore, designers have to find new ways to deal with big datasets.

Among all the possible methods available to analyze and communicate statistics, data visualization resembles to be the most powerful one (Edward Tufte, 2001)

Data visualization stands out for its capacity to simplify the given data in a condensed space (Krum, 2013) and minimize the loss of information (Fayyad, Wierse, & Grinstein, 2002).

Visualization enables us to see the entire data set with minimal eye movement without scrolling or flipping between pages. If we looked at a spreadsheet with 80,000 values instead, how long would it take us to get a general understanding of the market? (Krum, 2013)

3.4.6. INVENTIVE COMMUNICATION

Multiple academics have studied the positive influence of communication and social interaction on developing creative ideas (e.g. Gino et al., 2009; Farzaneh et al., 2012; Elisondo, 2016). In design, the creation process involves numerous stakeholders and departments; therefore, excellence in collaboration is almost a requisite for a creative result.

In inventive communication, data visualization is proven as one of the most valuable tools. Data Visualization offers a common language between designers, engineers, executives and data analysts to discuss and develop new ideas (Fiaz et al., 2016). Visuals could provide a platform for collaboration within organizations, speeding understanding insights and taking decisions.

3.4.7. PLAYFUL INTERACTIVITY

Data visualization is anchored on the importance of interactivity, notably to handle complex datasets (Steele and Iliinsky, 2010). As the expert Tamara Munzner (2014) has affirmed, the scope and capabilities of interactiveness have vastly increased thanks to computers.

This interaction allows the user to be active and engaged to detangle the different levels of insights, from understanding the overview to capturing small details. The reader has to transition between different visual displays, demanding his conscious participation. This feature of data visualization can have a great potential on creative sessions and could increase the designers' trust in the data collected, which often are skeptical if the data is not processed and interpreted by them (McGinley and Dong, 2011).

3.4.8. ITERATIVE PARALLELISM

Recent studies on the creativity field defend the cyclic and iterative nature of the creative process (e.g. Amabile, 1983; Tardif and Sternberg, 1988; Cross, 1997; Crilly, 2010). In parallel, most design process models coincide with the existence of creative stage loops. Considering these facts can be beneficial to use tools that support these cognitive cycles; in consequence, data visualization can be an ideal complement.

Data visualization allows exploratory processes that involve multiple rounds of data examination, visualization, reflection and deployment (McKenna et al. 2012).

4
REDEFINING THE SCOPE AND CHOOSING THE DESIGN DIRECTION

This chapter explains the final design direction selected to support Ford in using data visualizations as creative material. The first section presents the final problem statement synthesized from the research inside Ford. Subsequently, the design direction and design goal are described. Finally, the last section gathers the design requirements considered as meaningful for a suitable solution.

4.1. Problem Statement _ *Page 56*

4.2. Design direction & Design goal _ *Page 57*

4.1. PROBLEM STATEMENT

After having a general understanding of the current situation inside Ford and their challenges towards creativity and data visualization, the technique 5W1H was used (check Appendix). This technique stands for Who, What, Where, When, Why and How and is a checklist technique "of the most important questions to be asked to analyze a design problem" (Van Boeijen et al. 2014.) As a result, the following problem statement was defined:

> The Ford Design team wants to use data visualization to incorporate the data collected from its Smart vehicle concepts team into their creative process.
>
> In general, in order to use data, the Design Team needs to collaborate with the Global Analytics Data Insights department (GDIA). Due to privacy regulation and limited time-human resources, the GDIA only performs analysis to confirm or refute a defined hypothesis, commonly known as Confirmatory Analysis. As Explorative projects do not have a clear goal and are rooted in discovering opportunities, Ford Design Team usually struggles to propose these hypotheses or questions.
>
> Aiming to help the designers to create a specific demand for the GDIA, the CDF department has created a template with the specifications needed for any analysis, which can also be used in other design phases.
>
> Until now, the analysis results are delivered to the Design Team in different formats, mainly clean data sets or a Data Visualization to confirm or refute the hypothesis. The data is not explored further, denying the opportunity to obtain other interesting insights.

4.2. DESIGN DIRECTION & DESIGN GOAL

Considering the previous problem statement, from all the possible opportunities proposed in the chapter The intersection: Crossing bridges; "Exploratory inquiring" was revealed as the one with higher potential and most interesting for the company. This observed opportunity is based on the potential of Data Visualization to help generate questions that can promote insight finding and creative thinking, which mostly happens in Explorative data analysis, not Confirmatory Data Analysis.

Therefore, the ultimate design goal of this project is to

Support Ford Design Team to perform exploratory inquiring when facing data visualizations to facilitate their creative process.

» Exploratory inquiring: to iteratively generate questions that allow designers to find insights.

» Data visualizations: as seen in the chapter data visualization, there is a wide variety of data visualizations. Aiming to scope this project, I have selected the most used static charts used right now inside the team: histograms and scatter plots.

» Facilitate their creative process: obtain insights from the data that can fuel their knowledge and therefore, improve their problem framing and solution generation.

In order to undertake this design goal, this project mainly focused on the development of academic research. The goal is not necessarily to develop one concept but rather to offer possible design directions based on scientific research. The followings steps were considered necessary:

1. Understand the theoretical potential of Exploratory Inquiring (Chapter 5)

2. Validate and confirm the main principles of Exploratory Inquiring with empirical practices (Chapter 6)

On the other hand, some suggestions have been included to carry out new research and a guide to explore possible specific solutions that will support Ford in Exploratory Inquiring (Chapter 7).

5 EXPLORATORY INQUIRING

In previous chapter we have explored the data-enabled design process at Ford, delve into creative process, data visualization and the explorative data analysis and its interactions, and also selected the final design direction selected to support Ford in the use of data visualizations as creative material. Now is the time to know what exploratory inquiring is and why it is was considered the best option to the Ford Design Team.

5.1. Exploratory Inquiring _ Page 60

5.1. EXPLORATORY INQUIRING

The previous chapters explained that creativity involves knowledge, and questioning is one of humans' most potent reasoning mechanisms to gain that knowledge. As Berger (2014) indicates, "questioning enables us to organize our thinking about what we do not know". On the other hand, designers' work also relies on the questions they formulate (Eris, 2004), usually in an iterative process that drive them to discover insights that reframe their solution and problem space. On a data-enabled design approach, designers will have to perform this inquiry process while facing diverse data sources. However, which kind of questions does the designers' creative process demand? Furthermore, how can inquiring help them make sense of data? Finally, can visualization help them to engage in formulating that questions?

Based on data scientists, data visualization can be an invaluable tool, as it supports the exploration of datasets and the generation of more meaningful questions (Tukey, 1977; Fekete et al., 2012).

Building on these premises, I considered that using data visualization to generate questions and, therefore, discovering insights and improving the creative process (what I have named as Exploratory Inquiring in this research) is the most inspiring opportunity for the Ford design team. Exploratory Inquiring could help Ford's design team choose data visualizations based on the kinds of questions they can ask, or alternatively explore the kinds of questions they can ask based on the data visualization shown to them.

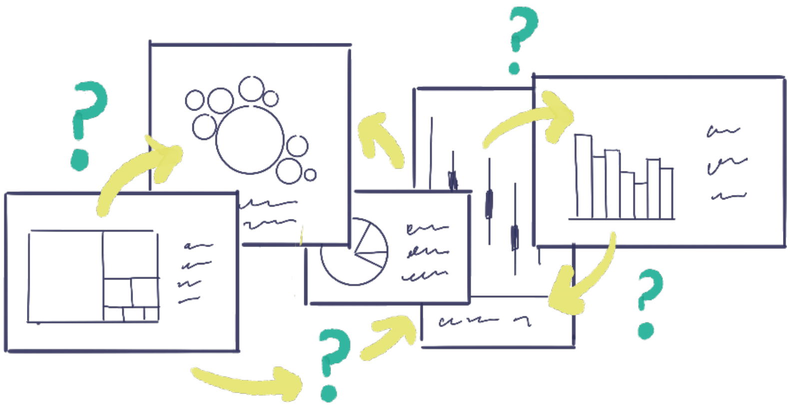


Fig 22. General idea of how the iterative process of Exploratory Inquiring could be.

The designer knowledge can be nourished with data through Exploratory Inquiring. In this process, when the designers are faced with a determined data visualization, their perception (P) will retrieve a certain knowledge input that will provoke the generation of questions. After this, the designer can choose to interact (E), determine new specifications (S) and visualize (V) a different data visualization. The iteration of this process can facilitate the generation of:

- » New research questions and hypotheses that might not be possible to have just from the initial dataset.
- » "Better" design hypotheses:
    - » **Deep reasoning questions** (Eris, 2004) that would enable the discovery of unexpected, complex, deep or relevant information (North, 2006)
    - » **Generative design questions** (Eris, 2004) that promote divergent thinking and, consequently, originate possible design solutions or ideas.

Based on the literature research, I expect that the kind of question asked by the designer is influenced by the kind of data visualization (Fig 23) and the designer's knowledge. Therefore, I considered understanding this influence relevant, and consequently, I executed different empirical studies gathered in the next chapter.

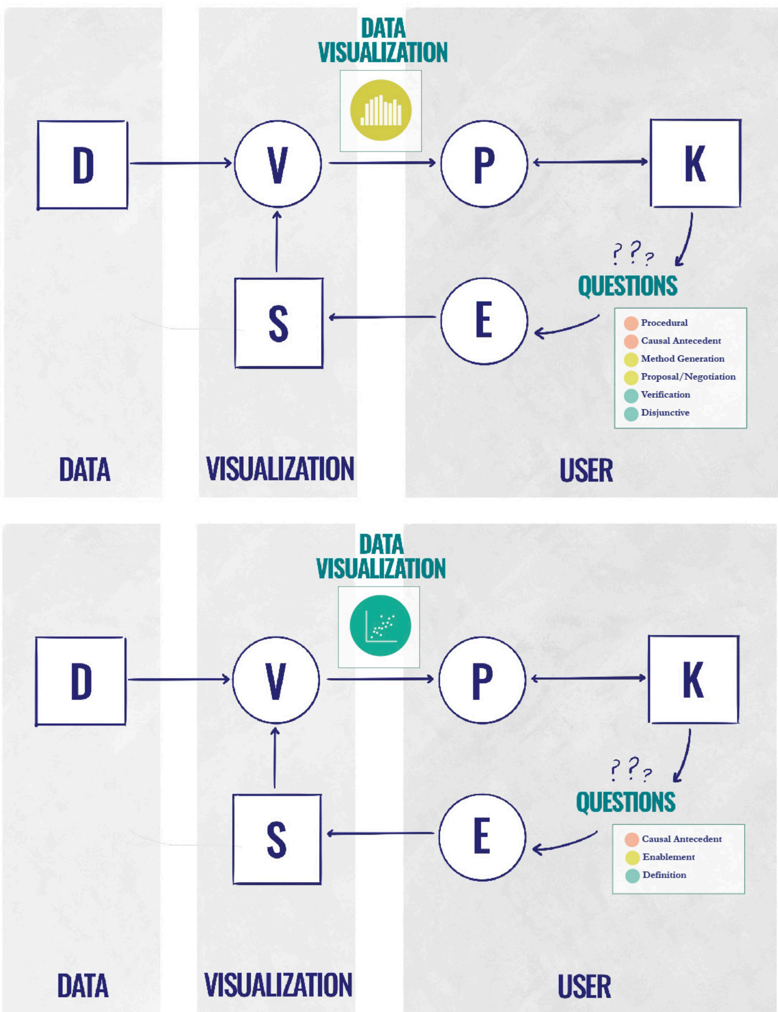


Fig 23. Based on the literature research, we could imagine that a determine type of data visualisation trigger specific types of questions in the designer

6 EMPIRICAL STUDIES

After reviewing the theory, understanding Exploratory Inquiring on a practical level was the next step. This chapter contains firstly an overview of the different studies, followed by the chosen analysis method. Subsequently, the description and results of each study are addressed. Finally, I discuss each study independently and in comparison, considering possible limitations.

6.1. Overview of the studies _ *Page 64*

6.2. Analysis Method _ *Page 65*

6.3. Study 1: Personal Explorations _ *Page 70*

6.4. Study 2: Students Workshop _ *Page 73*

6.5. Study 3: Workshop with Ford _ *Page 79*

6.6. Discussion _ *Page 84*

6.7. Limitations _ *Page 90*

6.1. OVERVIEW OF THE STUDIES

The empirical dimension of this project entails three studies (Table 5):

» The first one consisted of my personal exploration of the Exploratory Inquiring framework described in chapter 5. These practices aimed to understand the possibilities of exploring a dataset by constantly playing between question formulation and data visualization creation. Its full description and the most relevant observations can be found in section 6.3.

» Based on the preliminary findings from the personal explorations, a second study was designed to enable me to freely observe how a group of designers face data visualizations and which questions arise from them (described in section 6.4). The execution of this workshop led me to create an analysis method that, afterwards, I applied to all the empirical studies (in detail in section 6.2).

» Lastly, I replicated the workshop with the company employees (section 6.5) to explore how the level of expertise influences the generation of questions based on data visualizations.

Nº	EMPIRICAL STUDY	DATE	PARTICIPANTS
1	Students Workshop	14th April - 30th May	Me
2	Students Workshop	17th June	4 Design students from URP, Chair and Company mentor
3	Ford Workshop	11th October	7 Employees and Research assistant designer from URP

Table 5. Overview of the empirical studies conducted.

6.2. ANALYSIS METHOD

The analysis of the different empirical studies can be divided into two procedures:

» Observations and reflection analysis

The most interesting quotes and observations were selected in a mind map (see Fig 24 for reference). These findings helped in the discussion of the questions that emerged from the three studies.

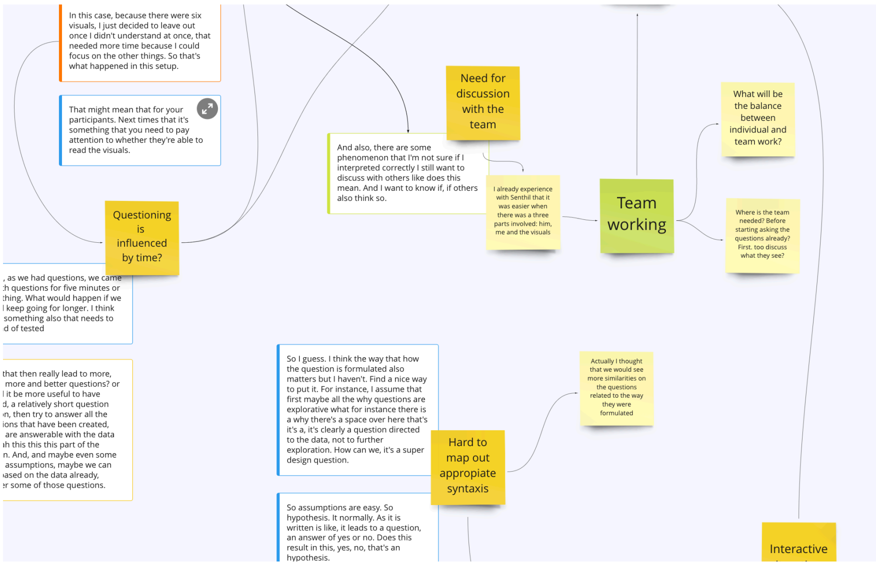


Fig 24. Example of the mind map developed with the participants quotes

» Questions analysis

Finding an appropriate way to analyse the relationships between the questions formulated and the visualisation was a tedious and iterative process. The final steps are described below in combination with an example (gathered at the end of this section, Fig 29 and Table 6):

1. Designate a question number. If the question was a follow up to one before, the same code was used with a "+" symbol. If the question triggered more than one direction of follow-up questions, a letter was also added (for instance, "1a+" and "1b+")

2. Identify the participant, taking into consideration their expertise on data visualisations and the database topic.

3. Note the type (distribution, correlation) and subtype (histogram, scatter-plot, connected scatter) of the data visualisation that triggered the question.

4. Identify the specific target: trend, feature and outlier (based on Tamara Munzner (2014) work). In addition, another category was added: Visual in general; because I could not determine which target triggered the question (Fig 25).

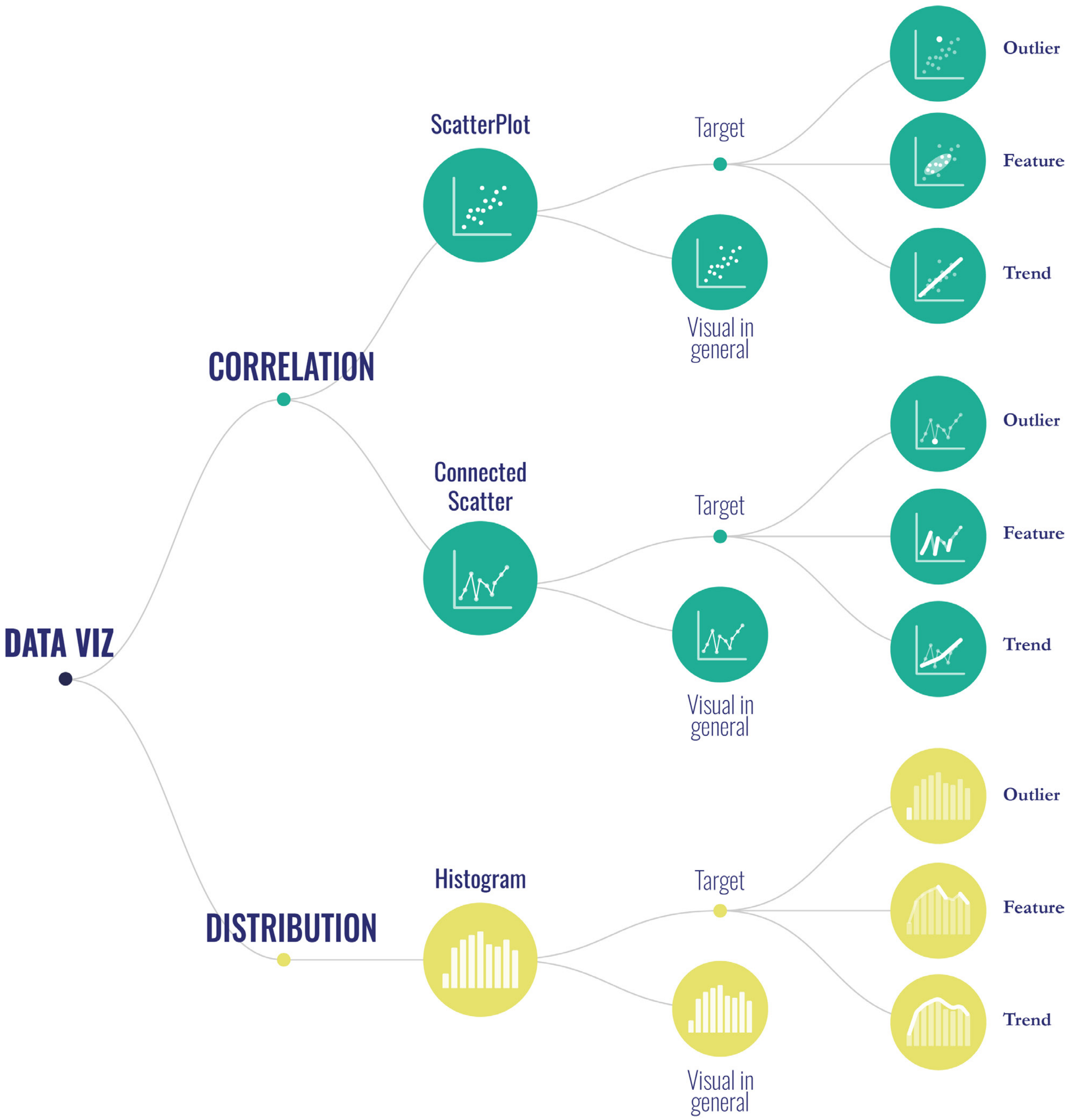


Fig 25. The coding scheme used for analysing the visuals that triggered the questions based on Munzner (2014).

5. When analysing the questions, I first focused on whether they reflected convergent, divergent thinking or any of them. Secondly, I categorised them based on Eris taxonomy (2004). During my analysis, I realised some questions could not be categorised strictly under one group (Low-level questions, Deep Reasoning Questions and Generative Design Questions). Therefore, I considered them a hybrid between two categories as suggested by Graesser, Rus and Cai (2008) (Fig 26). The final coding scheme is shown in Fig 27.

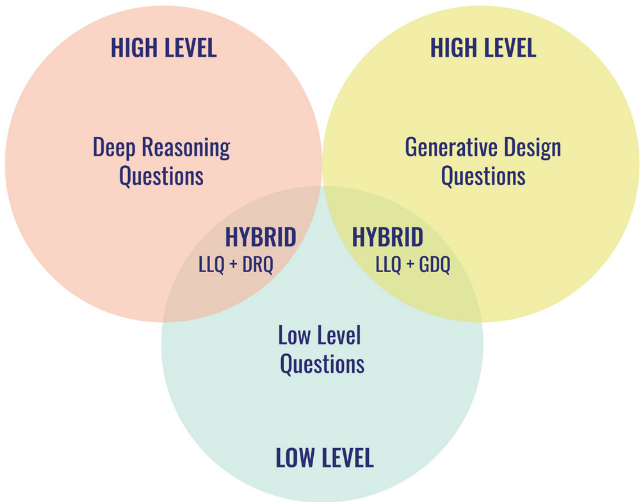


Fig 26. The hybrid questions resulted from a combination of a Low-level question with a High-level one.

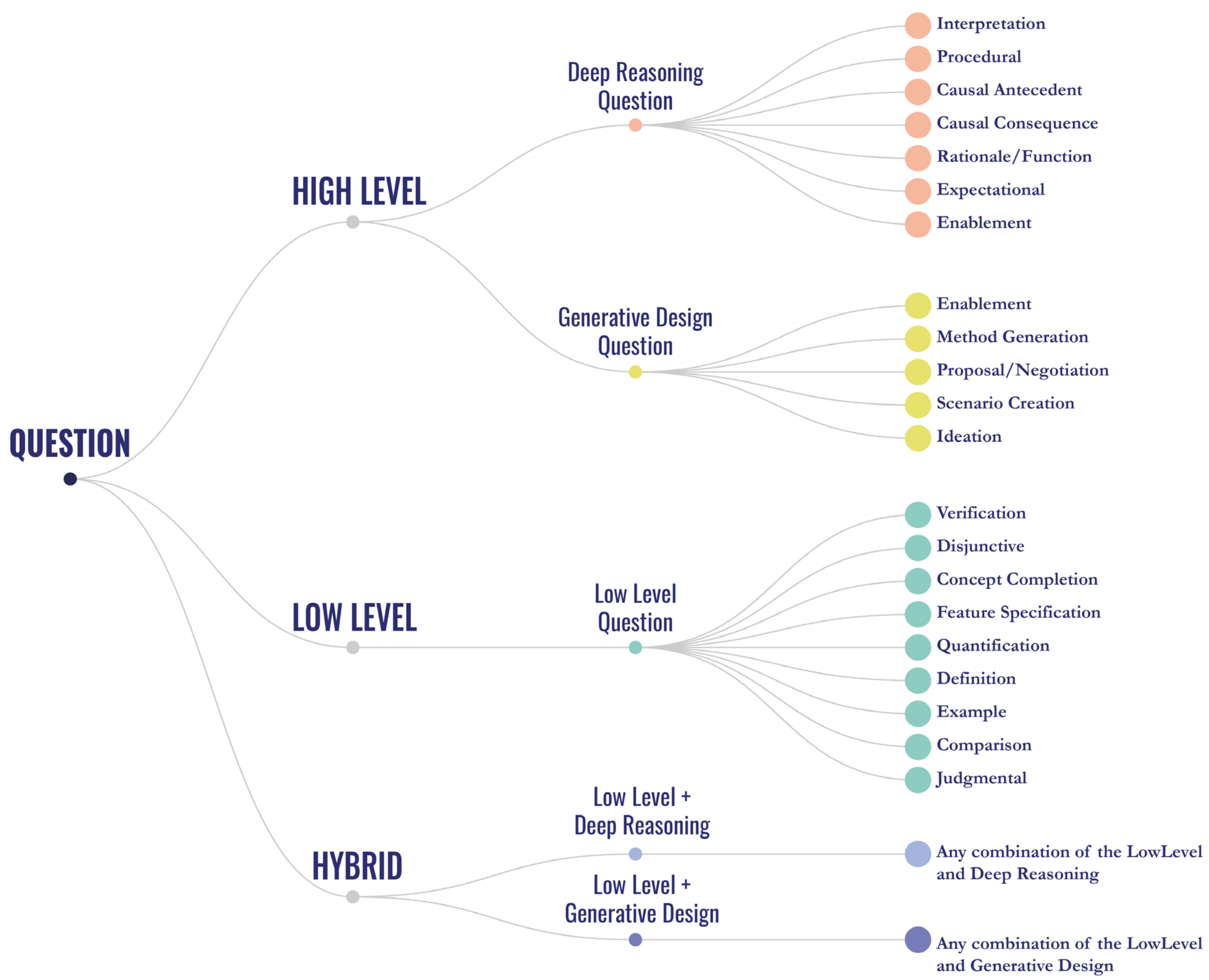


Fig 27. The final coding scheme used for identifying the type of questions extending the work of Eris (2004) by considering also hybrid questions as Graesser, Rus and Cai (2008)

6. Lastly, I created two dimensions of classification: implication and purpose, which were not covered in the founded taxonomies, which helped to understand the full context of the participants' questions. As Exploratory Inquiring is understood as an iterative process and not an isolated question, the first dimension determines the implications of the questions on the next possible steps for the designer . The second one, based on the participants' classifications, aims to identify the purpose behind the formulation of the question. The limitations of both categories are discussed in section 6.7.

Fig 28. Ford research illustration

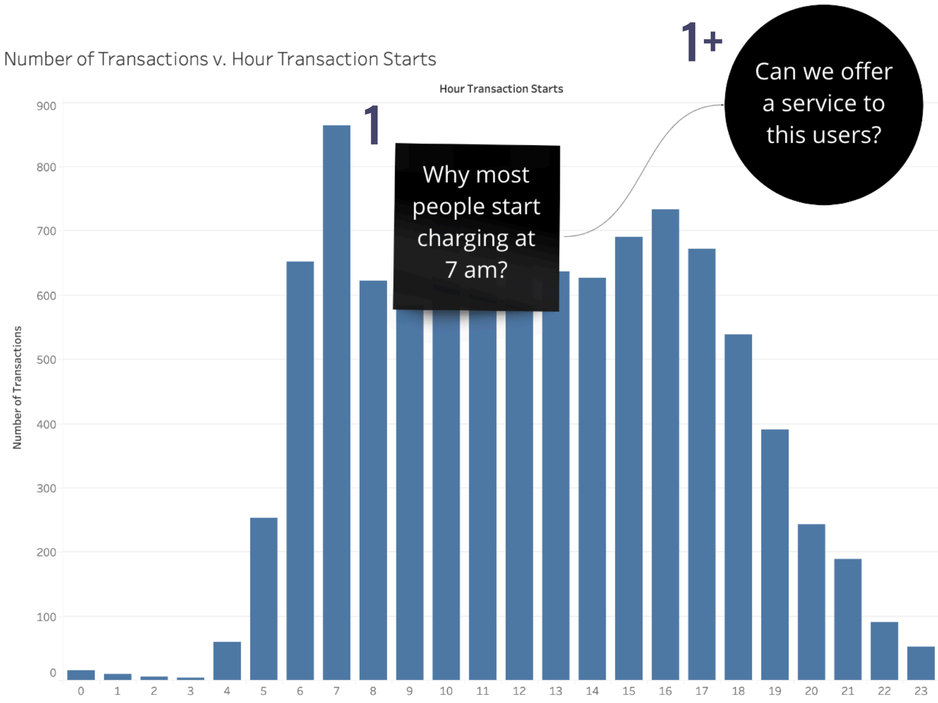


Fig 29. Example of one of the data visualisations used and the type of questions that could emerge from it, analysed in .

	Nº	**1**	**1+**
	QUESTION	*Why do most people start charging at 7pm?*	*Can we offer a service to these users?*
MUNZNER (2014)	**TYPE OF DV**	Distribution	Distribution
	SUBTYPE OF DV	Histogram	Histogram
	TARGET	Trend	Trend
ERIS (2004)	**QUESTION TYPE**	Deep Reasoning	Generative Design
	SUBTYPE	Causal antecedent	Proposal
PERSONAL CLASSIFICATIONS	**IMPLICATION**	Further research	Supports the solution space exploration
	PURPOSE	Identify types of users	Propose “way” to solution

Table 6. Example of how questions were analysed.

6.3. STUDY 1: PERSONAL EXPLORATIONS

6.3.1. OBJECTIVE

This first study was a self-practice of Exploratory Inquiring to understand further the findings reviewed in the literature research (chapter 3). Performing them could allow having a general overview of the possible challenges that Ford designers face when trying to use data visualisation as input for their creative process and confirm the possibilities of Exploratory Inquiring to support them.

6.3.2. SETTING

- **Duration:** The six sessions ranged from 1 to 3 hours.
- **Working Platform:** Miro, an online working environment similar to Blue Scape (the software used in Ford creative sessions)
- **Data Visualization Software:** Tableau and Voyager
- **Dataset:** Open dataset retrieved from the organisation ElaadNL[1], an initiative to improve the Vehicle Smart Charging experience developed by the Dutch electrical grid operators (Elaad, 2021). This dataset includes an overview of 10.000 random transactions from public charging stations operated by EVnetNL in the Netherlands.

Fig 30. Example of the personal explorations realized in a Miro board

6.3.3. PROCEDURE

The different sessions followed the Exploratory Data Analysis attitude reviewed in the literature research (3.3.4). Firstly, I visualised the dataset from Elaad with the software Tableau. Then, the data visualisation obtained was exported into Miro, and I registered the questions triggered. Finally, if I wanted or needed to explore further, I went back to Tableau to create new data visualisations. Thus, the process consisted of multiple iterative loops (see image Fig 30).

1 The dataset used for this studies can be found in https://platform.elaad.io/analyses.html

6.3.4. RESULTS[1]

The total number of questions that emerged was 33 (Fig 31). The majority of them were identified as Low-Level questions (36%). Deep Reasoning Questions and Generative Design questions had a similar rate, 18,18% and 24,24% respectively. A remarkable percentage of questions (21,21%) could not be categorized in one of these categories and were considered as hybrid questions. Bellow, the most noticeable results on each of these groups:

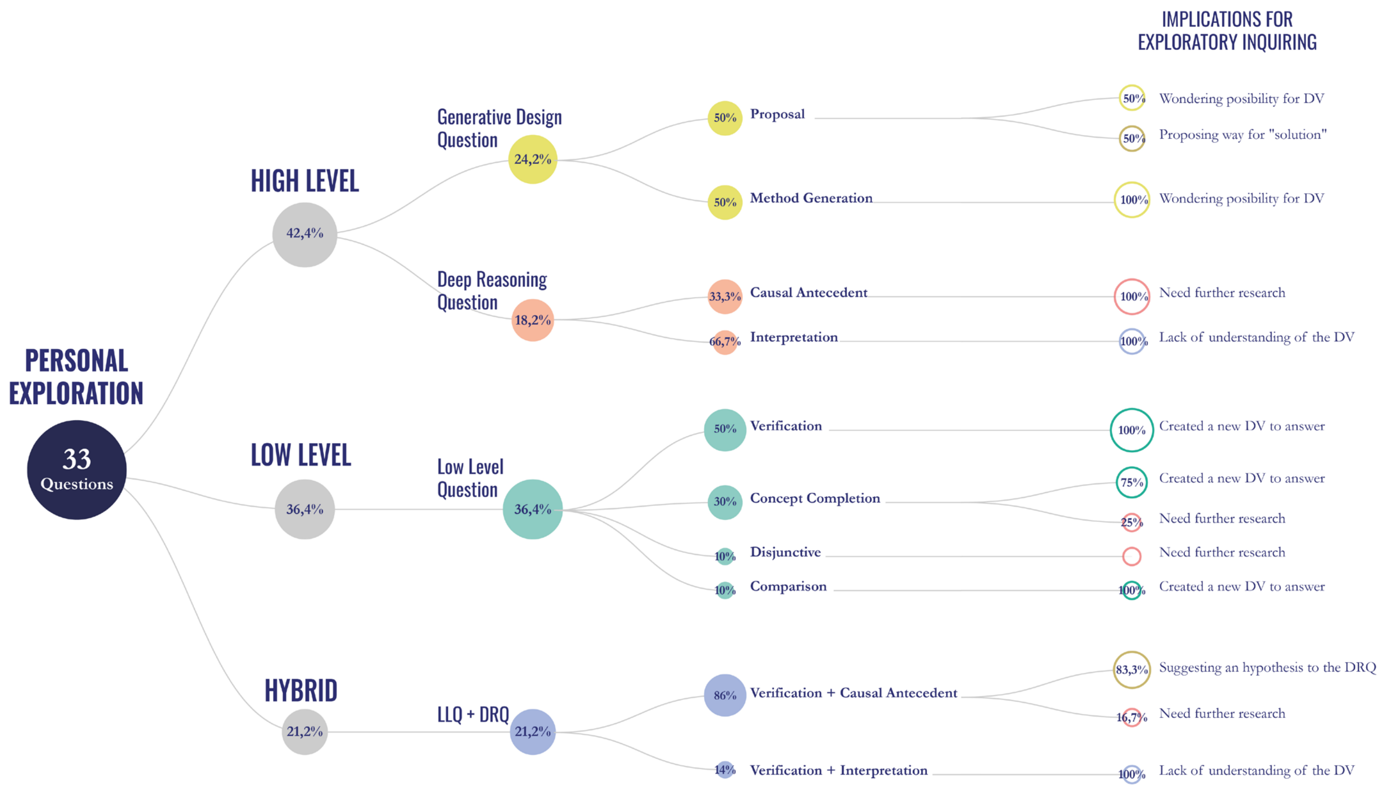


Fig 31. Question analysis results from study 1 (Personal Explorations)

- **Low-Level questions:** In most of the cases (83,3%), the formulation of these questions was followed by the interaction with the triggering visual (e.g.: zoom in or change the colour coding) or the creation of a new visual (sometimes also with other variables from the dataset). The rest of the questions needed further research to be answered. The subtypes of questions in order of incidence were: Verification (50%); Concept completion (30%); Disjunctive (10%) and Comparison (10%)

- **Deep Reasoning questions:** Only interpretation (33,33%) and causal antecedent (66,66%) questions were registered. The first ones were always a reflection of a lack of understanding of the visuals, the second ones could not be answered with the dataset available and open possibilities for further research.

1 The results and analysis of these questions can be found in Appendix: Personal Explorations

» **Generative Design Questions:** From the five possible categories, just two were used: Proposal (50%) and Method Generation (50%). A quarter of them was proposing an "idea" to help in the design of a better EV charging experience. The rest of the questions were formulated to explore the possibilities for creating new visuals, as a tool to diverge and generate potential alternatives to explore further the dataset.

» **Hybrid questions:** in the personal explorations, I observed a combination of Low-Level and Deep Reasoning Questions. Specifically, always a Verification together with an Interpretation (14%) or a Causal Antecedent (86%) (the only subcategories registered isolated on the Deep reasoning group). The Verification + Interpretation, was a reflection on the lack of understanding of the data visualization, trying to verify if what I was interpreting was real or if there was a problem in the data visualization. ("Actually there are really few returning customers in this group no?"). The rest of the questions suggest a hypothesis when asking about the cause of something, on two occasions this led to the creation of a new visual while the majority would need further research.

In relation to the type of visuals created, the participants were more inspired by the histograms (Fig 32). Therefore, most of the questions were also triggered by this subtype of the distribution graph. There is a predominance on the target, being 54% elaborated on features (in comparison with 15% outliers and 6% trends); afterwards, the target was not explicitly stated and were considered as visual in general (24%)).

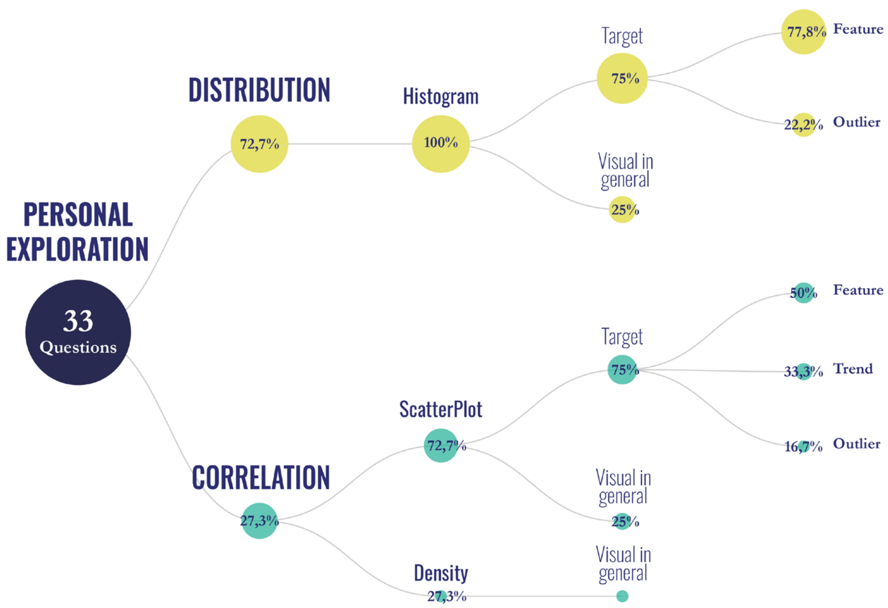


Fig 32. Question and Data Visualization analysis results from study 1 (Personal Explorations)

6.4. STUDY 2: STUDENTS WORKSHOP

6.4.1. OBJECTIVE

This was an experimental study to gather insights about how designers intuitively approach exploratory questioning with data visualizations, to look for correlations between the data visualization and the questions triggered and to observe possible similarities and differences between the participants.

6.4.2. RESEARCH APPROACH

The workshop pretended to mimic a real situation where a design team started an Explorative Project at Ford. Based on their design thinking process (Section 2.5) the workshop could be an activity of the phase Gather research & inspiration (Fig 10). In this particular case, a possible design problem is given to the team along with data visualizations to explore to obtain inputs for their creative process.

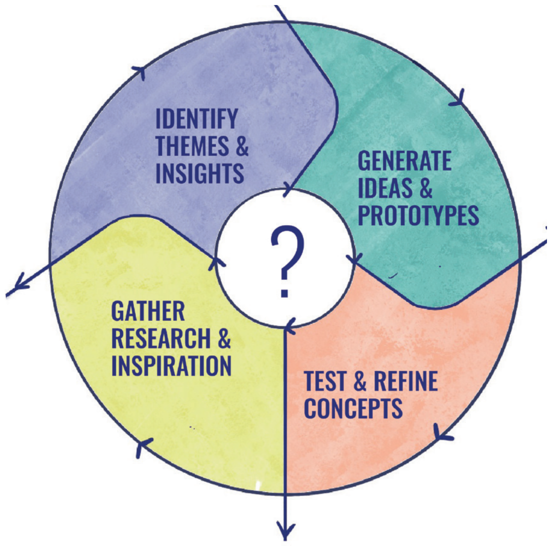


Fig 33. Ford research illustration

6.4.3. SETTING

- » **Date:** 17th June
- » **Duration:** 45 minutes
- » **Platform:** Miro, an online working environment similar to Blue Scape (the software used in Ford creative sessions)
- » **Participants:** Mentors and leaders of the Ford & TU Delft URP: Nicole Eikelenberg and Milene Gonçalves + 4 design students also graduating in the Ford URP for Data-enabled Design. All participants knew the context of the project but were not aware of the coding structure or the actual goal of the workshop. Only one participant, Nicole Eikelenberg, designer at Ford, can be considered experienced both in the dataset topic and in data visualizations in general.
- » **Dataset:** Taking into consideration the privacy limitations described in the section 2.6, this study did not use Ford's data although the findings

could be applied to their data. The dataset used was the same as in the Personal Explorations, namely the ElaadNL one. The dimensions used in this study are listed in the Table 7 below along with the description provided by the organization and an example of the raw data.

	Description from Elaad (2021)	Data Example
TRANSACTION ID	The unique transaction code	3491779
CONNECTOR ID	Many charging stations have two connections (two sockets for charge plugs) and this indicates what connector was used for the transactions.	1
UTC TRANSACTION START	The moment the transaction was started (logged in locale time zone).	27/8/2019 14:52:00
UTC TRANSACTION STOP	The moment the plug was disconnected and the transaction was stopped.	27/8/2019 17:58:19
START CARD	The RFID card (hashed) which has been used to start a transaction.	0c24de2f8216313f75daf876ec-7c2223e17c866462ae41ca8d-9c98a30b222ac1
CONNECTED TIME	Time difference between the start and end of a transaction.	3,11 h
CHARGE TIME	Total time wherein energy transfer took place	3,1 h
TOTAL ENERGY	The total energy demand (kWh) per session.	9,86 kWh

Table 7. Data dimensions and example used in this study from Elaad (2021)

» **Data Visualizations:** A total of six visualizations were created for this study. All belong to two types of static charts: histograms and scatter plots. This decision was taken based on the following reasons:

» Limiting the data visualization possibilities to be able to make deeper observations.

» This type of graphics are currently within the most used by the SVC team, therefore could be more beneficial for the team to understand better how to use them.

» During my personal explorations (described in section 6.3); I relied mostly on these two styles. By choosing the same type for the workshops, comparing behaviours would be more accurate.

6.4.4. PROCEDURE

The workshop followed the structure as indicated in the following Fig 34:

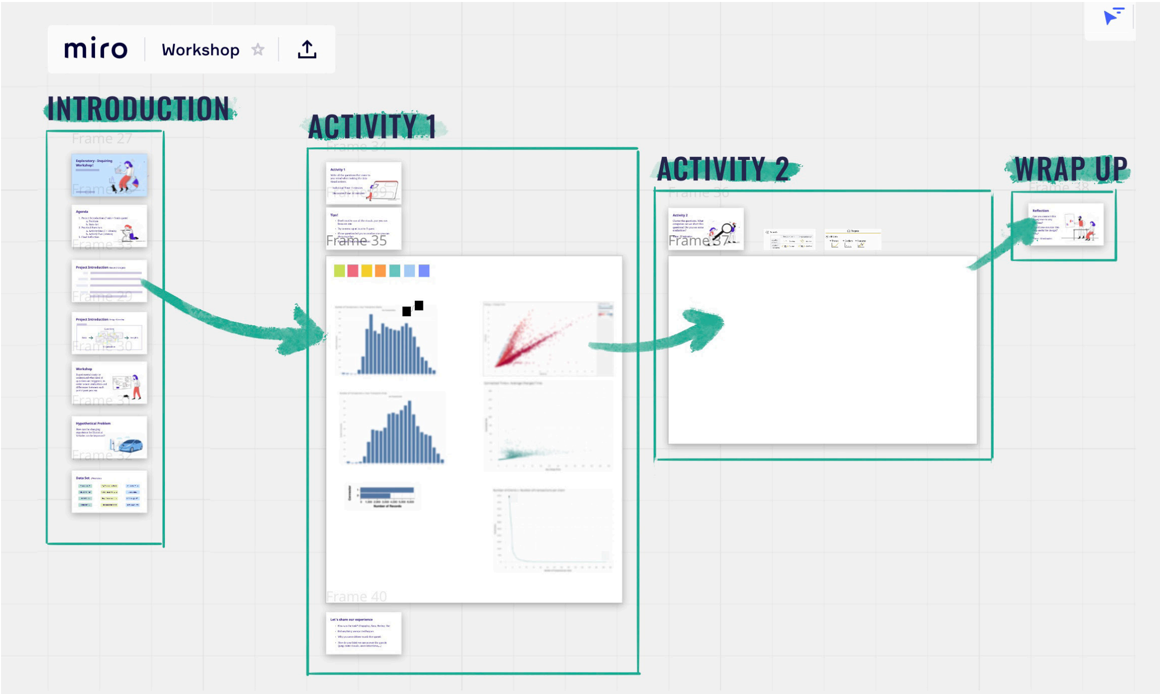


Fig 34. Miro boards used in the study 2

- **Project introduction** (10 min). First, the main insights retrieved from the research were presented to the participants to introduce the goal of the session (explore how designers formulate questions when faced with data visualizations in a design project). Subsequently, the hypothetical design problem was given: *How can the charging experience for Electrical Vehicles be improved?* The dataset from Elaad.nl and its dimensions were described. After the presentation, the participants were invited to ask if they had any doubt.

- **Activity 1** (15 min). Participants were asked to individually write in digital post-its all the questions that came to their minds when looking at the graphs for 5 minutes. Before the activity, the following tips were given to set a more intuitive, open and relaxed mindset:
  - *Don't need to use all the visuals, you can just focus on one*
  - *Try to come up with at least 3 questions*
  - *If one question led you to another one you can link them together*
  - *Have fun! You are just exploring :)*

  The activity was followed up with a group discussion to share the experience.

- **Activity 2** (10 min). A collaborative exercise to cluster the questions. During the exercise, participants were asked to deliberate together on the similarities and differences of the questions.

- **Wrap Up** (10 min). The last minutes were dedicated to reflecting together on the experience of Exploratory Questioning with Data Visuals.

6.4.5. RESULTS[1]

At the end of the workshop, 33 questions were registered from Activity 1 (Fig 35). The majority of the questions were a hybrid combination of Low Level Questions and Deep Reasoning (36,36%). The second more used questions were Deep Reasoning Questions (32,7%). Generative Design questions and Low-Level questions have an incidence rate of 18,8% and 12,12% respectively.

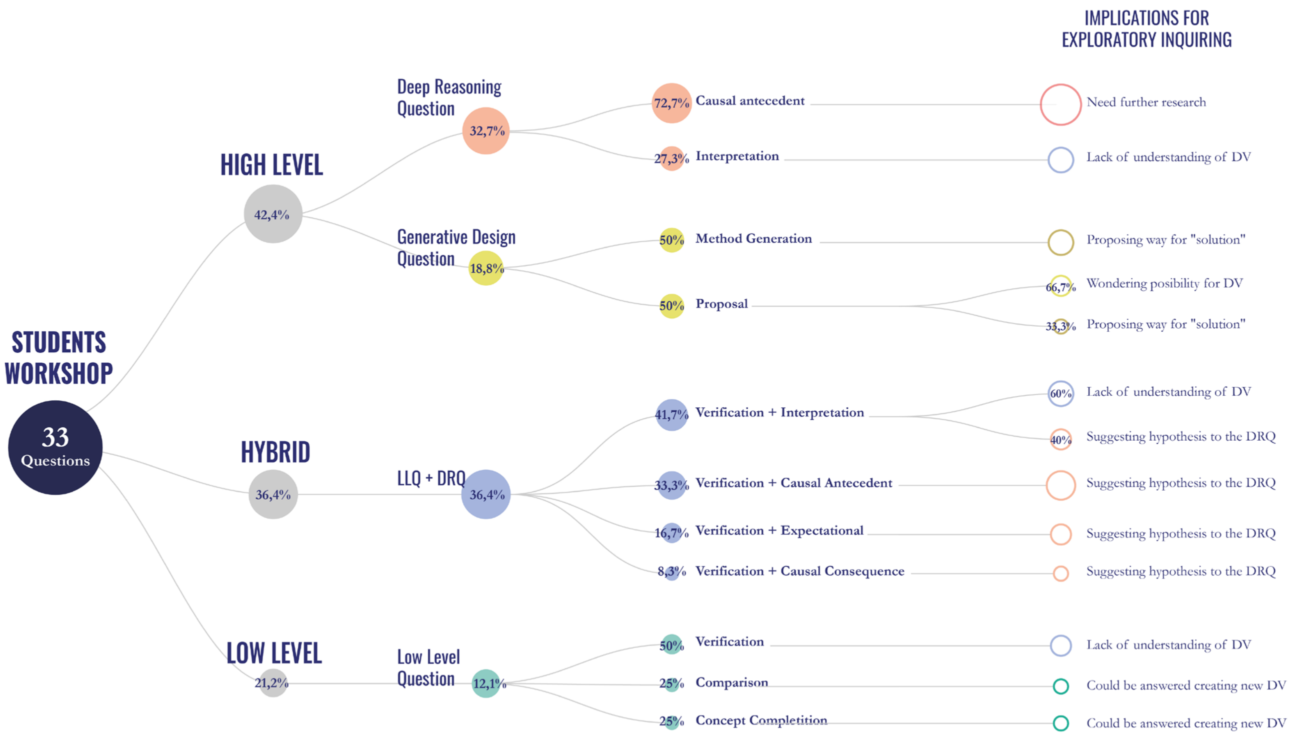


Fig 35. Question analysis results from study 2 (Students Workshop)

- **Low-Level questions:** There were identified three subtypes of questions: Verification (50%); Concept completion (25%); and Comparison (25%). Half of these questions were formulated due to the lack of understanding of the data visualization while the other half could be answered with a new data visualization.
- **Deep Reasoning questions:** The participants formulated Interpretation questions (27,27%) when they were not understanding the data visualization. The other subtype of questions registered was causal antecedent (72,73%), in this case, to answer them further research will be needed.
- **Generative Design Questions:** only two categories were used: Proposal (50%), inspired by correlation graphs, and Method Generation (50%), resulting from distribution graphs. In general, these questions served to propose an idea to distribute equally the charging events (66,67%). Only the ones formulated by the expert in the topic domain (EV charging), also more experienced than the rest of the participants on the use of data visualizations, suggest possibilities for creating new data visualizations to understand the dataset better.

1 The results and analysis of these questions can be found in Appendix: Students Workshop

» **Hybrid questions:** The questions registered were always a combination of a Verification (Low-level questions) with a Deep Reasoning question, specifically: Interpretation (41,66%), Causal Antecedent (33,33%), Expectational (16,67%) and Casual Consequence (8,33%). In the majority of them (75%), the participants were wondering about the cause of something they observed on the graphs and suggesting a possible hypothesis that could be the reason (75%). The rest of the questions (25%), respond to a lack of understanding of the data visual but still suggests a possible interpretation of what they perceive, being always a combination of Verification + Interpretation.

The participants were more triggered by the histograms in comparison to the scatter plots, 58% versus 42% (Fig 36). Despite the slight difference, the most remarkable fact is that almost half of the questions registered in the scatter plots were related to the lack of understanding of the visuals itself (43%) and only one question was hybrid and suggested a hypothesis (7%) (formulated by the expert participant). Specific targets such as outliers (33%) and features (30%) seem more inspiring than trends (19%) and the visual in general (9%).

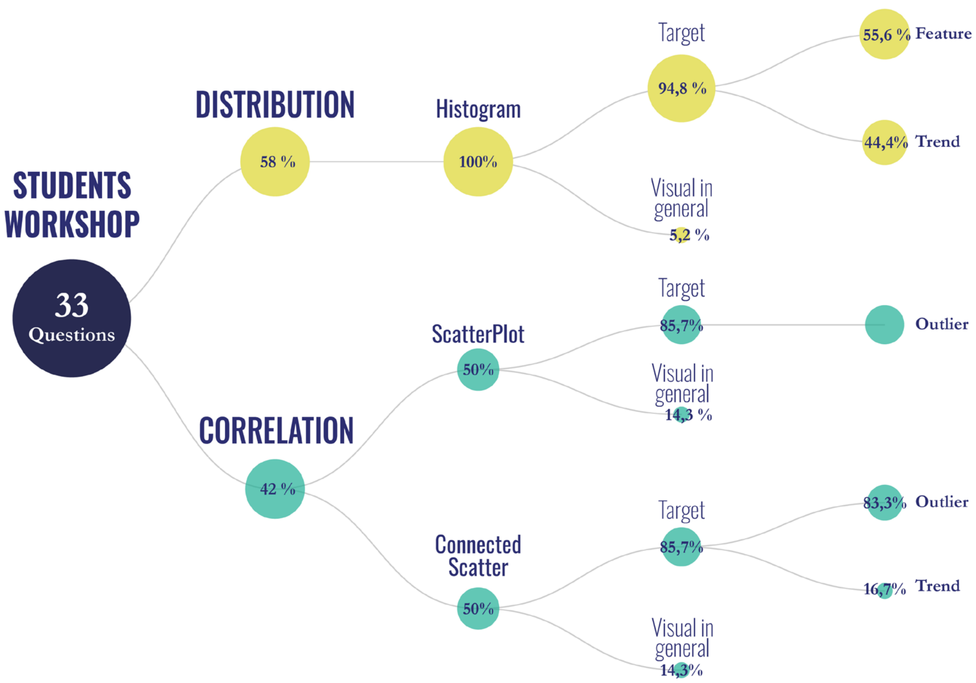


Fig 36. Question and Data Visualization analysis results from study 2 (Students workshop)

In relation to the classification done by the participants in Activity 2, the participants created four categories:

» **Questions that lead to ideation:** the participants clearly identify the Generative Design Questions that proposed a way to a possible solution, not considering the two GDQ formulated by the expert to create new data visualizations.

» **Verify assumptions:** all the questions identified in this group, except question nº 13+ (*And to which factors could that be related?*) are hybrid questions, hence a mix between Verification and a Deep Reasoning Question.

» **Questions that will need further research:** the participants considered that most of the questions require further research (36,3%), in this group we found a mix of Deep Reasoning (66,67%), Generative Design (8,34%) and Hybrid questions (25%).

» **Questions directed to the data visuals:** only one question in this group does not seem to be strictly related to the data visualization (nº 29: *What kept this user coming here over and over again*). The rest, I also relate them with the data visualizations either because of the lack of understanding (50%), the proposal of a new graph (12,5%) or the possibility to be answered with another visual (25%).

The reflection after Activity 2, extended in time (approx. 32 min), was recorded and automatically transcribed in the software Otter. Overall, these are the most remarkable comments:

» The experience was considered positive. The main potential mentioned by the participants in the possible follow up research resulting from this experience. The workshop opened their possibilities and in their own words, at some point:

> ***"there should be a convergence moment, so that we can move forward"*** (Participant 3, 2021).

» Participants clearly exposed that they felt limited by their knowledge on data visualization to formulate questions.

> ***"But my ability to ask deeper questions really depends on whether I understand the graph or not"*** (Participant 3, 2021)

> ***"In this case, because there were six visuals, I just decided to leave out the ones I didn't understand at once."*** (Participant 2, 2021)

» Another interesting annotation is that participants claimed the need to contrast information with other team members, especially to clarify doubts about the visuals.

> ***"I'm not sure if I interpreted correctly the visuals and I still want to discuss with others and I want to know if others also think so".*** (Participant S, 2021)

» Participants were surprised when classifying the questions because it was effortless to identify ideation questions, but they struggled to create other classification groups. As a reason for this difficulty, they suggested the variety of language composition and words use (like interrogative adverbs)

6.5. STUDY 3: WORKSHOP WITH FORD

6.5.1. OBJECTIVE

As said in the discussion from the students' workshop, the main observation was the difference between the expertise participant and the novice ones. This contrast led me to execute a final workshop (Study 3) with the company employees with the intention of exploring the influence of expertise in the questions formulated.

6.5.2. RESEARCH APPROACH

The workshop replicated the approach of the students' workshop (check 6.4.2)

6.5.3. SETTING

Below are the setting details that changed in this study with respect to the first one:

» **Date:** 11th October

» **Duration:** 1 hour 30 minutes

» **Participants:** Seven Ford employees with different backgrounds: 2 participants from the D-Ford organisation, 2 participants that are experts in EV charging and 3 members from the Smart Vehicle Concepts team. The eighth participant was a URP research assistant with experience as a designer and designing with data.

6.5.4. PROCEDURE

The workshop was adapted to Ford, below I detail the modifications and the reasons for these changes:

» **Project introduction** (20 min). In the first workshop, the participants belonged to the graduation community of students at Ford and they were already introduced to my topic. In this second workshop, I modified the introduction slides to offer a more complete overview of my project as the participants did not know anything before.

» **Activity 0** (15 min). Additional activity to warm up. The participants had to choose a post-it colour and write their experience on the topic of EV charging and on data visualizations. When most of them were done, we had an introductory round.

- **Activity 1** (20 min). The duration of the time given to formulating questions was extended from 5 to 8 minutes, as it seemed short for the students in the previous workshop.
- **Wrap Up** (25 min). I decided to increase the time of the final reflection based on the total length taken during the first workshop.

6.5.5. RESULTS[1]

In this case, the participants filled up 62 different post-its, among them, there are some particular cases:

- Two questions (nº: 1 & 2) were formulated before the first activity was fully explained and the visuals were shown, therefore they are considered as not valid.
- Six comments[2] with the participants' statements on their interpretations (nº 44: *It seems that most people are connected longer than the vehicle is charging*).
- Two notes (nº23 & 56), although they include a question mark, I interpreted them as comments. The punctuation symbol seems to be used with the purpose of hedging their words (nª23 *I have no idea what this means?*)
- Four of the collected annotations were not written as a direct question (including a question mark) but as comments (nº17, nº24, nº28, nº31). No matter the semantic formulation, all of them seem to be GDQ, specifically, proposal questions. As an example, nº17, *Would like to see how the end-time correlates to the time people start the transaction*, can be understood as: *Could we see if the end time correlates to the time people start the transaction?*

Taking into account these exceptions, there are 52 questions (Fig 37). Both hybrid questions (28,85%) and Low-Level questions (28,55%) record the highest incidence. Deep Reasoning Questions and Generative Design questions have the same percentage (21,15%). The most outstanding results are presented below:

- **Low-Level questions:** The subtypes of questions identified were : Disjunctive (26,67), Feature Specification (26,67%), Quantification (20%), Verification (13,33%), and Concept completion (13,33%). The majority was formulated due to the lack of understanding of the visuals (53,33%), followed by questions that need further research (40%); only one of the questions could be for sure answered with the creation of a new data visualization.
- **Deep Reasoning questions:** the Causal Antecedent questions are more frequent (45,45%), after that Interpretation (27,27%), casual

1 The results and analysis of these questions can be found in Appendix: Workshop with Ford
2 Considering comment as "a statement that expresses an opinion about something" v. a question "a sentence, phrase, or word that asks for information" (Cambridge, 2021)

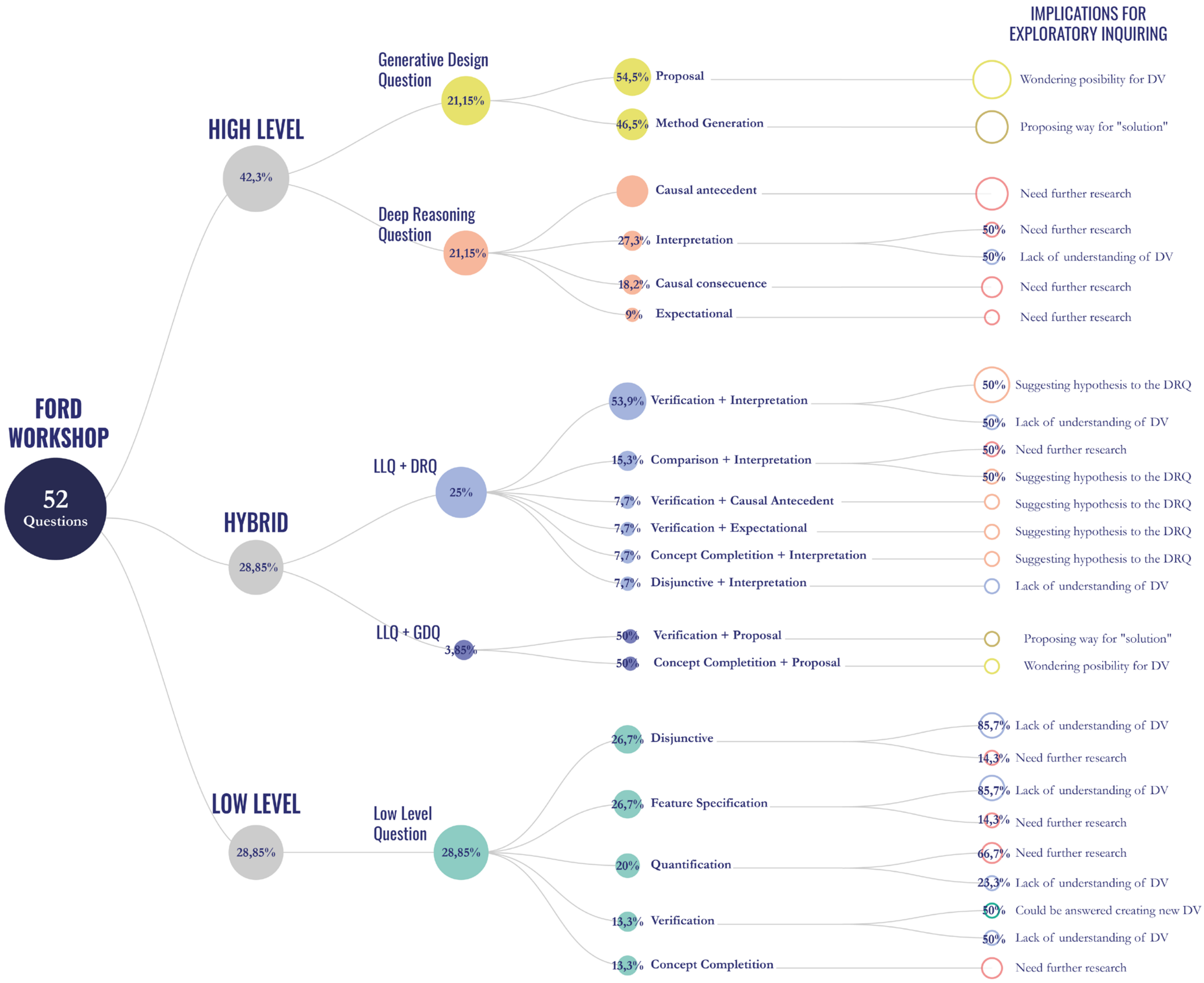


Fig 37. Question analysis results from study 3 (Workshop with Ford)

consequent (18,18%) and Expectational (9%). Interpretation questions correspond to the participant's purpose of understanding better the data visualization, the rest would probably need further research to be answered.

» **Generative Design Questions:** Generative Design Questions: in this workshop, the same behaviour is repeated, there are only records in two categories Proposal (54.55%) and Generation of Method (45.45%). In terms of intention, each of them was clearly used with an intention: suggesting the creation of new visuals (Proposal) and enabling the creation of ideas to equally distribute the users (Method Generation).

» **Hybrid questions:** on this occasion, there were not only Low-Level and Deep Reasoning Questions but also LLQ with Generative Design ones.

» **LLQ + DRQ:** Sometimes Verification is joined to Interpretation (53,85%) , Expectational (7,69%), or Causal Antecedent (7,69%) and in others Interpretation come together with Disjunctive (7,69%) or Concept Completition (7,69%) or Comparison (15,38%). Is hard to identify a pattern that correlates the type of questions and the intention of the participant. In general terms, they were used to suggest a possible hypothesis (77%).

» **LLQ + GDQ**: In a small percentage (3,85%), Proposal questions are observed attached to Verification (50%) or Concept Completition (50%).

The participants focused mostly on the correlation graphs (63,5%) versus the distribution ones (Fig 38). While it is true that almost a third of the questions triggered by the scatter plots were due to the lack of understanding, and in comparison, only 10% of this type of question arose from looking at the histograms. Finally, it is remarkable that more than half of the questions refer to the visual in general (52%). Features were the preferred target (23%), followed by outliers (13%) and trends (12%).

Ford employees categorized the questions into eight categories: charging behaviour, customer, creating new visuals, related geographical data, questions about the visuals, new areas of exploration and questions about the energy

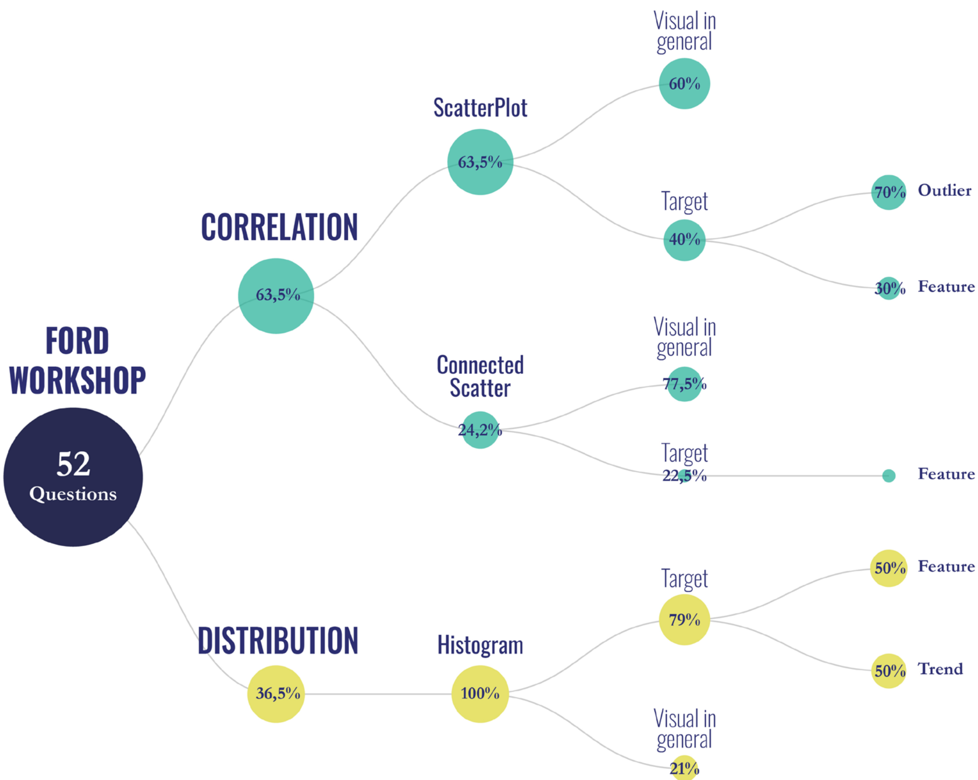


Fig 38. Question and Data Visualization analysis results from study 3 (Workshop with Ford)

system.

Regarding the final reflection, a few observations are noteworthy:

» Most of the participants focused on explaining their discoveries about the topic visualized rather than the general experience.

» They appeared engaged in the activity (Fig 39).

» The employees also looked beyond the workshop and speculated about future possibilities that I took into consideration in section 6.7 (Fig 40).

Actually that I did not expect during like the working hours, which, which I find quite interesting.

Which, which I find quite interesting.

I think that's what I was always looking for was like extra bits of information to bring to this.

Yeah, thought provoking session.

Fig 39. Quotes from the participants which reflect their interest on the session

It would also be interesting, also a next step to see, what can we learn from this data and how can we combine it with the data we have from our Ford vehicles.

So in these visualizations we don't have any information on the vehicle side (...) but that's information that we could, I assume, get from our data.

The experience definitely generated a lot of questions and I am wondering what would be the questions that we would ask knowing which extra data was available.

And just thinking about that it would also be great to create a stakeholder map, and then ask the questions from each of these different perspectives, because I think also all these perspectives will raise different types of questions.

Fig 40. Quotes from the participants proposing future possibilities to incorporate this activity.

6.6. DISCUSSION

6.6.1. DISCUSSION PERSONAL EXPLORATIONS

D.1 Until analysing these explorative practices, I did not know that I was essentially using histograms. I believe this predisposition happened because univariate summaries are one of the most basic forms of analysis as they allow exploring each variable separately. Therefore, they are usually considered one of the first steps to explore a dataset (Wongsuphasawat et al. 2017). Still, even if I just relied almost on one type of data visualization, I could find a way to answer the Low-level questions in an 83,3 %, either by changing the encoding in the data visualisation software or by creating a new visualisation. These accomplishments could lead to speculation if, with higher skills in data visualisation, I could also have answered higher or hybrid level questions. For instance, looking back in time, now that I am more experience in Data visualisation, I could have responded to the question nº15 Is because in the stations was really few transactions registered? by revising the dataset with other tools.

D.2 Another observation from my personal experience is how easily I was feeling lost and losing track of the path. Performing an Exploratory Data Analysis can be overwhelming; however, it can be facilitated with a clear system to organize findings, visualize the questions and the steps followed which agrees with the observations from NLI researchers, revised in section 3.3.4, (Gotz et al. 2006; Shrinivasan & van Wijk, 2008; Stasko et al., 2008; Gotz & Zhou, 2009; Ragan et al. 2015).

6.6.2. DISCUSSION STUDENTS WORKSHOP

D.3 As expected the most used questions were Deep Reasoning (69,1%), either isolated (32,7%) or in combination in a hybrid question (36,4%), as the participants wanted to understand the reasons behind the data representations. Based on Eris, the questions on which designers rely most during a conceptualization phase (reviewed in section 3.2.5) are Causal Antecedent questions. This is verified in this experiment, as almost 37% of the total questions, either isolated or in a hybrid combination, belonged to this group. However, other reasoning mechanisms that directly addressed causality are Expectational and Causal Consequence, which are not observed by Eris in that phase and in this case, have also a high incidence. This could be explained because Eris' conceptualization experiments did not take into account a context where designers use external sources like data visualizations where they need to interpret what they perceive. In order to do an interpretation, the viewer or reader will need to ask not only why something happened (Causal Antecedent) but also What it means (Interpretation), why something did not happen (Expectational), and what the effects were (Causal Consequence). This can highlight the possibilities to further determine the questions observed by Eris (2004) for each design stage.

D.4 The difference in questions between the distribution and correlation graphs might be the result of two main causes: the limited time of the session and the level of knowledge both in the topic and in data visualizations of the participants. The participants focused on making questions on the histograms which in their opinion were simpler and more intuitive. Apart from the confirmation of the verbal reflections, we can observe that the questions triggered by the correlation graphs are mainly a reaction to their lack of understanding of the visual.

D.5 Almost a third of the questions could be answered with another data visual or by explaining the missing information related to the lack of data visualizations' knowledge. This observation can reaffirm the interest in performing Exploratory Inquiring, hence iterating between both questions and data visualizations. On the other hand, this led me to believe that the time designers expend to formulate questions due to the lack of understanding could have been minutes used to generate questions about EV charging. Therefore, it could also be essential to create more "readable" graphs, work together with the data analysts or give educational support on data visualization.

D.6 Generative Design Questions were the least frequently asked questions, which is consistent with the theory of Dym et al. (2005). It is also confirmed that the most common question subtype is Proposal/Negotiation, as in the studies by Eris (2004). It is observed that according to their intentionality they can be divided into two types:

» Generating new ideas to modify the distribution of users in the charts. These questions were easily identified by the participants as Ideation questions. All of them were asked by the students.

» Proposing the creation of new graphs to obtain another perspective of the database. In this case, there were only two questions, both of them from the automotive expert participant, not only more experienced on the topic but also in data visualization. This observation could mean that the background knowledge highly influences the possibilities for continuing the exploration of questions and the dataset.

D.7 As the majority of the questions are specific targets (outliers and features), this sparked the question of whether there is a predisposition on the way designers observe the data visualization (look at the details or have an overview first) or if it is could be influenced by the knowledge background of the designer.

6.6.3. DISCUSSION FORD WORKSHOP

D.8 Before discussing the type of questions, let me now turn your attention to the participants' comments. Although the instructions asked for questions, the participants noted down their interpretations of the graphs as facts. These comments could have been written as Verification + Interpretation hybrid questions (for instance, nº57 *This graph only indicates that the majority of people have charged around two times*. could be understood as: *Does this graph mean that the majority of the people have charged just two times?*) but it seems that the employees' confidence and/or expertise led them to declare it as clear statements. Nevertheless, they still wrote them so maybe they were surprised or not expecting what they "read" in the visualization.

D.9 Other interesting comments are the replies between participants (see detail below Fig 41). Furthermore, in the afterwards reflection, the *light-blue post-it* employee remarked the appreciation to that interaction. This discovering sharing moment and the positive response to it, made me think about which benefits could have to formulate the questions as a team (discussed further in section 6.6 with the other experiences).

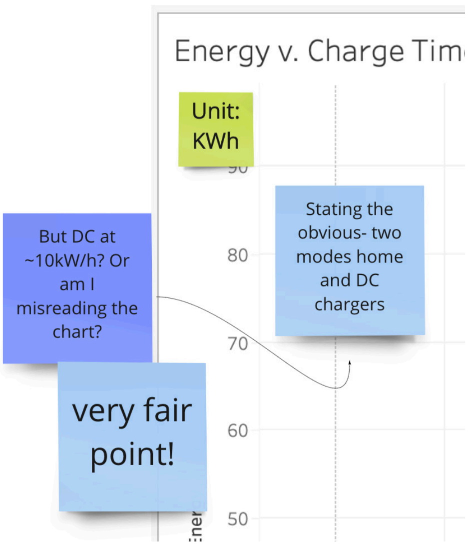


Fig 41. Conversation between participants during the workshop

D.10 Lastly, there were comments that appeared to be questions and vice-versa. While nº23 & 56 included a question mark but were statements, nº17, nº24, nº28, nº31 looked like comments but could be considered proposal questions (GDQ). These findings could be relevant to reconsider the rubric of the taxonomies to fit possible language variations especially.

D.11 In relation to the Ford employees' questions, Low-level questions were the most prevalent either isolated or in combination in a Hybrid question. However, a lot of these isolated questions can be potentially considered as a hybrid or a follow-up to a Deep Reasoning question. Employees ask for specific details (*What was the layout of the chargers?*) as they have knowledge on the topic but probably as the result of asking firstly to themselves a Deep Reasoning question (*Why did customers only charge once during this time interval?*).

D.12 In both workshops, there is almost the same questions' ratio due to a lack of understanding. Nevertheless, in this case, the Low-level questions are formulated to ask precisely for the missing details, like the units in graph 4. In comparison, the students formulated more general questions that express their lack of data visualizations' knowledge.

D.13 In this workshop, we can see a 7% increment in the questions that propose the creation of a new data visualization. All of them were formulated by an employee with medium-high experience in data visualization.

6.6.4. COMPARISON OF THE THREE EXPERIENCES

The following diagram (Fig 42) considers all the subtypes of questions used by the participants in each study, either alone or in combination in a hybrid question:

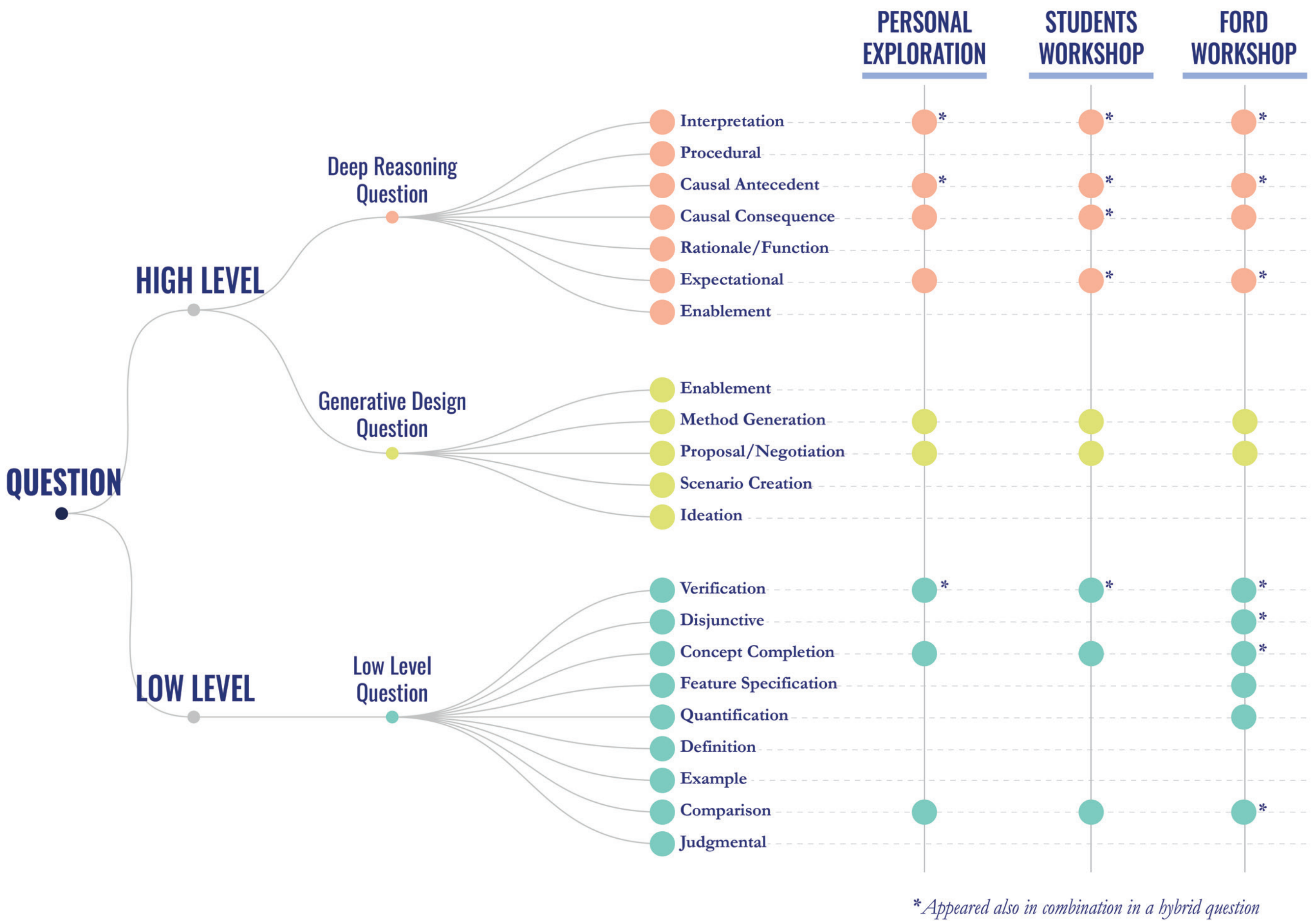


Fig 42. All the question types observed in the different studies.

D.14 In the three studies, participants always used the same kinds of Deep Reasoning and Generative Design Questions. However, we can see an increase of Low-level subtype questions in Ford's workshop. Instead of just simple Verification questions, the employees formulate questions like Feature specification, Concept completion or Quantification. On the one hand, these results on the Low-level questions could agree with the observation made by Cross (2004): novices do not know which precise information to search for, while more experienced designers do have. Furthermore, academics have observed how more experienced designers can address the information taking into account previous experiences from the past, like possible solutions (Lawsoon, 2004; Petre, 2004). Perhaps this is why Ford employees, instead of general questions, asked for specific characteristics (for example, inspired by the third visual, *What was the layout of the chargers?*).

D.15 The subtype of Generative Design questions used was the same: Method generation and Proposal. Nevertheless, there was a clear distinction in the intention. Students realise more questions focused on ideation for the solution. At the same time, the domain experts and I preferred to formulate GDQ to propose DV possibilities that will lead to a deeper exploration of the dataset. One possible explanation, related to the co-evolution model, is that maybe design students are more solution-focused while the Ford experts tend to tackle design problems in a problem-focused way (Lawson, 1979; 2004; Lloyd and Scott, 1994). Furthermore, this emphasises that the background knowledge on data visualisation influences the possibilities of further exploration and coming out with specifications for the new visual.

D.16 After the three studies, I populated the following matrix to see the possible influence of the data visualisation type (distribution, correlation). At this point, I believe that it would be possible to observe almost any kind of question derived from any visual. However, the relevance is finding which ones have a more substantial influence and which ones are more intuitive for designers to formulate. For example, on the high-level questions, the subcategories observed were the same in the three studies, and I could not detect in any of them a Procedural (DRQ), Rationale/Function (DRQ), Enablement (DRQ, GDQ), Scenario Creation (GDQ) or Ideation (GDQ). However, looking at the first graph given to the participants (Number of transactions v Hour Transaction starts), we could formulate a question on the non-registered categories: *What if there is an extra benefit offered during the night?* (Scenario Creation). Therefore, I would suggest performing a study with a more significant population and considering only the question types with higher incidence rates.

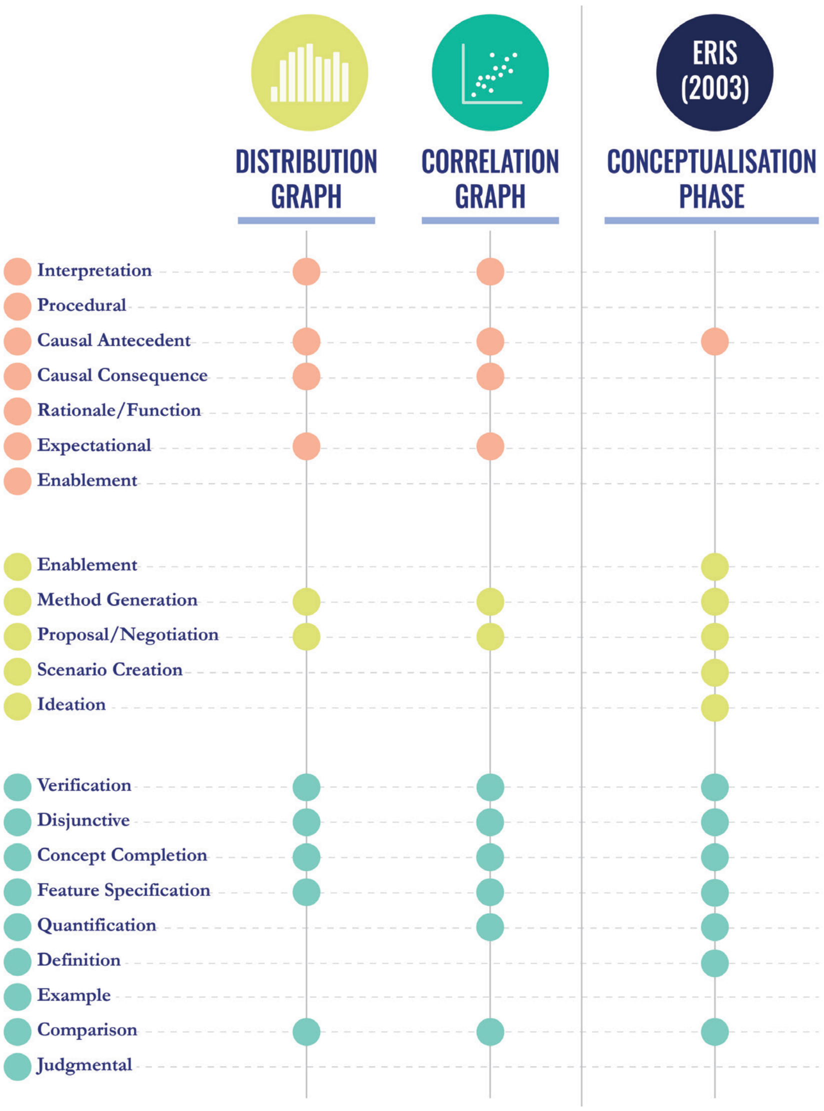


Fig 43. Comparison between the questions which Eris (2004) observed in the conceptualisation phase and the results obtained in this project depending on the type of visual

D.17 Eris (2004) highlighted that designers rely on particular questions' types (LLQ, DRQ, GDQ) depending on the design stage. As said in section 3.2.5, from the phases defined by Eris, Conceptualisation is the one matching our context: "involves users-needs finding, requirements definition, and idea generation". Fig 43 confronts the questions types observed by Eris and the ones noticed in this project. As also discussed in 6.6.2, participants relied on other Deep Reasoning questions (Interpretation, Causal Antecedent, Expectational). A possible explanation is that Eris did not consider the use of external sources like interviews, videos, or, in this case, data visualisations. Therefore, he did not find applicable rates in his statistical distribution of these questions' types, which were predominant in this project's context. As participants, in order to do an interpretation, they needed to ask not only *Why something happened?* (Causal Antecedent) but also *What does it mean?* (Interpretation) *Why this did not happen?* (Expectational) *What were the effects?* (Causal Consequence)

D.18 In the last workshop (study 3), a significantly higher number of questions were registered even with the brief amount of increased time. Of course, their level of expertise could have facilitated their formulation of questions, but also, we can consider the level of engagement. By this, I mean that even if the students were working under a problem statement, they already knew that we would not arrive at a design phase. Instead, Ford employees are already involved in the automotive sector even if they do not belong to a design department..

D.19 Surprisingly, there is a big difference between which visuals inspired the questions. Ford employees formulated more questions in the correlation graphs while both the students and myself focused on the distributions graphs (Fig 44). The correlation graphs from this study tackled specific variables of EV charging, which could be the barrier that made novices focus on the histograms; meanwhile, Ford experts were more engaged on that "new/particular" perspective. Besides, another variable affected the preference for a visual: the experience that participants have with data visualisation; as mentioned by the students, they were more used to histograms and therefore, these graphs were more intuitive for them.

Fig 44. The employees from Ford were more inspired by the correlation graphs and formulated more questions about them.

D.20 Also, we can notice a distinction in the way of reading the visuals; students attention was driven to specific targets, while for the majority of Ford employees' questions, it was hard to identify the inspiring target, and therefore categorised as "visual in general". One possible reason is the tactic knowledge of the experts, which makes their experience harder to articulate as they do not explicitly display their practice (Nightingale, 2009; Cross, 1982).

D.21 Some students stated the need to share interpretations with other members, while Ford participants already took the chance to communicate with comments between each other. In my particular case, I also observed the benefits of examining the dataset with others.

D.22 Lastly, we find a high percentage of questions in the three studies due to the lack of understanding of the data visualisations. However, verbalising these questions can also spark possibilities for solving them by, for instance, providing the designer with another set of visuals. This situation was observed in the first study where I, as both designer and data visualiser, could iterate on the visualisations to solve my lack of understanding.

6.7. LIMITATIONS

L.1 It is likely that the timing influences the nature and frequency of the questions formulated. In these studies, the time interval for question generation was relatively short and therefore, in extended time circumstances, the observations may vary.

L.2 In some cases, the intentions of the participants' questions and the inspiring target of the data visualization are implicit (especially with the domain experts as discussed, possibly because of their tacit knowledge) and subject to interpretation. These factors make it undoubtedly challenging to identify the type of question in the chosen analysis method.

L.3 As a collective exercise, the possible influence of the questions asked by some participants on others is undeniable, which can compromise the results of the studies. However, since designers tend to work in teams, it is also true that it could be a more realistic environment.

L.4 As a consequence of how the studies were conducted, the order of formulation of the questions is unknown, which does not allow us to analyse whether a predefined question pattern can facilitate an Exploratory Inquiring process.

L.5 The wording of some questions makes it difficult to include them in one type or another, so we must consider the bias of the categorizer, in this case, myself.

L.6 The context of the design activity, as a simulation, could only resemble the context of a design project in the industry. Furthermore, a difference in engagement was observed between the students and the employees.

7 CONCLUSIONS

This section contains recommendations for future improvements and their implementation, some of them are linked to the discussion points [D] or the limitations [L] for a better understanding. In any case, as has already been said in this document, this field still has great potential to be explored.

7.1. Exploratory Inquiring Implementation _ *Page 94*

7.2. Future recommendations _ *Page 96*

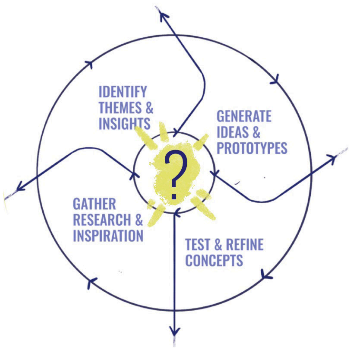
7.1. EXPLORATORY INQUIRING IMPLEMENTATION

The empirical studies suggest the potential already seen in the literature reviewed to use data visualization for Exploratory Inquiring and therefore support the designer's creative process at Ford.

Firstly, using Exploratory Inquiring can allow the Smart Vehicles Concept team to generate questions about the data. Verbalizing their questions increases the opportunities to find interesting questions to explore. Furthermore, being aware of the type of questions can bring some light on what kinds of visuals designers might need or can be used in the opposite way knowing which kind of data visualizations can help answer that specific type of questions.

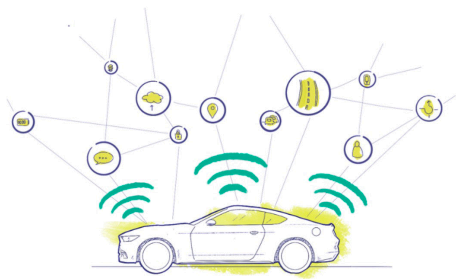
In order to show the potential that Exploratory Inquiring could have for Ford, I have created the following scenario, including the possible action paths.

1



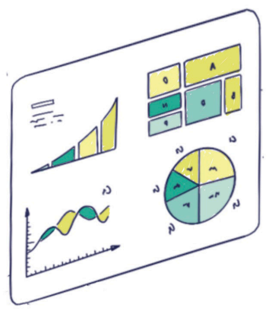
Start with the central question of the design thinking Ford's cycle

2



Select a dataset that could be interesting to explore

3



Create a data visualization to start exploring!

4

Look at it .... What do you see? What questions arise?

5

Analyze your question, what does it tell you? Based on your question what can you do?

6A

Improving the knowledge on data visualization can increase your opportunities as a designer to explore the dataset and discover more insights

6B

Can you answer the question by interacting with the visual? or Do you require to create a new visual to get the answer you need?

6C

Use this question opportunity to explore possible solutions to the problem you are working on!

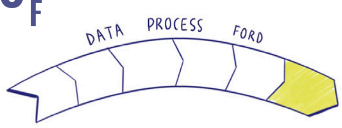
6D

Sometimes the questions arisen can be answered with other variables from the dataset that you have not visualize yet.

6E

Perhaps you have find an interesting topic to further research but you need to refine the question. Take a look on the questions guide to inspire you.

6F



Combining other data sources can complement your discoveries, e.g.: user interviews or other vehicle sources.

7.2. FUTURE RECOMMENDATIONS

7.1.1. PROPOSALS FOR THE LIMITATIONS

» A possible way to avoid misinterpretations on the questions intentions or which part of the visual was inspiring could be to conduct a follow-up interview with the participant to discuss these topics personally. For example, in Ford's workshop, a discussion with the participant could have clarified if some Low-Level questions might be considered the following of a Deep Reasoning question that the participant did not explicitly write. [L.2]

» Another interesting area to improve is the way to document the formulation order of the questions. Maybe then we could understand if a generative design question precedes a Low-level question or other patterns. In this case, I believe video recording the exercise could be the key to observing this topic which was not feasible for me in my studies with so many participants. Otherwise, if replicating the same experience, participants could be asked to number their annotations. [L.4]

» To be sure of accurate classification of the questions, a categorization performed by multiple researchers can be considered. [L.5]

» To overcome time and context limitations, an experiment could be done in the industry for full validation. [L.1][L.6]

7.1.2. OTHER POINTS TO CONSIDER IN THE FUTURE

» **Define the possibilities when iterating between questions & visuals** [D.1] [D.5] [D.9] [D.22]

One of the key points considered for implementing Exploratory Inquiring is understanding which action the designer can take after formulating the question. The diagram created to guide Ford is based on the literature research and my personal explorations (study 1). Consequently, further refinement could be done by explicitly validating each of the actions.

How can the action plan of a type of question be defined?

» **Communication & Collaboration** [D.21]

Through the studies, which explored the individual creation of questions, the participants noted the need to share and discuss the information. Furthermore, in my personal experience sharing my discoveries about the dataset with my supervisors helped me reformulate new and more inspiring questions. These observations can let us consider that an experience that encloses individual and group work on formulating questions could be more beneficial.

In education, multiple studies have observed how students' learning process improves when developing questions in groups.

Furthermore, one of the benefits of data visualisation is that it offers a common language between designers, engineers, executives or data analysts (Fiaz et al., 2016). Therefore it could also be interesting to incorporate both perspectives (designers and data scientists) in the experience. For example, performing an activity similar to the one proposed between the SVc team and the GDIA could help them understand each other.

» **Support designers Data Visualization level** [D.5] [D.6] [D.13]

As mentioned in this thesis, designers are not usually trained to work with big data or data visualisations (Davenport et al., 2019; Jung et al., 2019; Kun et al., 2020; Data-Centric Design-Lab, 2020). Likewise, in the empirical studies, it was observed how the knowledge in data visualisation could restrict the participants' questions. For instance, no novice came out with a high-level question to suggest a possibility for a different visualisation, while the more experienced on the field did.

Maybe a more in-depth introduction to familiarise themselves with the data set and the data visuals could play a high role in the subsequent questions from participants.

Perhaps if designers are more aware of the possibilities of DV, they could themselves figure out the specifications they might require from the data scientist when developing the data visualisations.

What if designers have a higher level of data visualizations? What if they can have more resources to explore the dataset?

» **Improving the Data Visualizations for designers** [D.12]

Considering that many of the questions were due to the lack of understanding of the data visualisations, it might be interesting to analyse if there are common patterns. Therefore we can extract possible requirements the data visualisation or dashboard should address to improve the readability or understandability by designers.

Think about what the designers require, what they need? How could the Data visualisations be more accessible for them?

» **Specific question generation activities!** [D.6] [D.14] [D.15]

The development of this project emphasises the importance of asking better questions for the designers as other researchers did (Eris, 2003; Dym et al. 2005; Eris, Sheppard & Kwan; 2007; Ahmed & Aurissicchio, 2007; Grebici et al. l., 2009; Aurisicchio et al. 2010, Cardoso et al. 2014, Cardoso et al. 2020, Hurst et al. 2021). Formulating questions can help in the creative process in multiple ways, like assessing the data more critically or framing the design problem. Moreover, the different taxonomies revised describe how each type of question also led to a different cognitive mechanism. Therefore, future studies could explore if designers could benefit from exposing themselves to the creation of specific types of questions, similar to a brainstorming session. For instance, in the Generations of ideas % prototypes phase, Ford could consider using data visualisations to inspire the formulation of just Generative Design Questions, contributing to exploring the solution space.

What if designers have a session where they can just use diverging questions?

» **Perception perspectives**

The present research emphasised the importance of background knowledge when formulating questions. Besides, the perception of each visual is different from each individual. Considering Ford's interest in applying a human-centred design approach, it could be interesting to "read" the visuals as if they were their users, asking the questions from the different stakeholder perspectives.

What would be to read these data visualisations with other eyes? What would our vehicles' users will see here?

» **Observing the influence of knowledge** [D.14]

New research could study differences between novice and experienced users and repeat the same process with updated data to confirm or refute if pre-existing knowledge or knowledge gathered through the process can improve it.

How does exploring a dataset change through time? How do designers expertise influence?

» **Refining the question**

One of the possibilities not addressed in this project was to develop the initial questions of the participants further. Using data visualisations could be the starting point for generating questions and then categorising, prioritising and reframing the most valuable questions, creating a higher value questioning technique.

How can the initial questions be improved? how can we choose the questions to further analyse?

» **Facilitate the process of Exploratory Inquiring** [D.2]

Future research could focus on providing the right tools to facilitate the process of exploratory inquiring. For example, by developing a clear system for organising findings and supporting the user to verbalise their questions explicitly, both aspects are recognised as valuable in exploratory data analysis software (Gotz et al. 2006; Shrinivasan & van Wijk, 2008; Stasko et al., 2008). Hence, designers could document better, be more aware, and engage with their own process exploring the dataset (Gotz & Zhou, 2009; Ragan et al., 2015).

Which tool can support designers to formulate questions when looking at a data visualization?

» **Guiding the designers** [D.7] [D.20]

The data collected suggested that students were more inspired by specific targets like outliers or features. By contrast, analysts suggest starting with a broad overview to familiarise themselves with the data and later, to concentrate on specific details (Moore and McCabe, 1989; Shneiderman, 1996). Specific data visualisation exploratory tools like Voyager 2 are also based on these principles (Wongsuphasawat, 2017). Future research might investigate if suggesting a particular order to read the visual could improve the designers' formulation of the questions.

Would it be better to offer a guideline or rules in the order of looking at the visual?

» **Explore other data visualizations and the questions triggered**

In this project, we observed that both histograms and connected scatter plots triggered the same types of high-level questions in the three studies, while a greater variety was observed in low-level questions.

It would be interesting to confirm or refute these initial findings, investigate other types of visuals and also not only to adress which type of questions data visualisations inspired but specifically if a particular visualisation can answer a determined type of questions.

Can a specific visual be more effective than others to answer a specific type of question?

» **Expand & Improve the taxonomies** [D.3] [D.8] [D.10] [D.11] [D.16] [D.17]

The possibility of refining the existing taxonomies warrants further investigation:

» **Interdisciplinary focus**

The taxonomies used for this project are based on the designer's perspective, which, as reviewed in the literature, generally has an abductive approach compared to the deductive approach characteristic from the data analyst (Jansen, 2020; Siqi, 2021). Hence, revising the taxonomies with the perspective of a data scientist could support the designers in data-enabled design approaches.

Table 8 gives an example of combining the designer and the data scientist perspectives to illustrate this possibility further. The first one is based on Eris (2004) and the second one on Choi et al. (2019) study on Natural Language Interfaces, which translate the user question to the data visualisation in a logic-based language.

What if designers and data scientists find a translation between their questions?

Table 8. Simple example of this possible way to incorporate a more interdisciplinary approach on the taxonomies

TYPE OF QUESTION	DESIGNER	DATA SCIENTIST
Verification + Interpretation	Do these people come to charge while they are working?	Does [Working time] correlate with [Starting charging time]?

» **Language variations**

Providing further examples on each categorisation could be valuable in analysing and categorising the questions and as a more inspiring guideline for designers (Already started in Appendix: Eris Taxonomy completed with Personal Examples). Specifically, data-enabled design approaches could benefit by addressing specific examples related to data.

Are words more inspiring for designers? Could designers be inspired if having more possibilities to express their questions?

» **Design phase**

This project manifested that Eris categorisation based on the design phases can be further refined. As explained, Eris did not consider external stimuli as interviews, data visualisation or others in the Conceptualisation phase, which are the main element in data-enabled design approaches. Therefore, I propose to revise this design phase and split it at least into two stages: gathering research and generating ideas (taking Ford's design thinking cyclic process as inspiration)

What types of questions can designers rely on the most when interpreting data visualizations?

8 REFERENCES

Ackoff, R. L. (1989). *From data to wisdom.* Journal of applied systems analysis, 16(1), 3-9.

Ahmed, S., & Aurisicchio, M. (2007). *A Model of Designing: Understanding Engineering Design Activity.* In Connected International Conference on Design Education. The University of New South Wales Sydney.

Amabile, T. M. (1983). *The social psychology of creativity: A componential conceptualization.* Journal of personality and social psychology, 45(2), 357.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). *Assessing the work environment for creativity.* Academy of management journal, 39(5), 1154-1184.

APA. (n.d.). *APA Dictionary of Psychology.* Confirmatory Data Analysis Definition. Retrieved September 28, 2021

Aurisano, J., Kumar, A., Gonzalez, A., Leigh, J., DiEugenio, B., & Johnson, A. (2016). *Articulate2: Toward a conversational interface for visual data exploration.* In IEEE Visualization.

Aurisicchio, M., (2005). *Characterising Information Acquisition in Engineering Design.* (Doctoral dissertation). Cambridge University.

Aurisicchio, M., Ahmed, S., & Wallace, K. M. (2007, January). *Improving design understanding by analyzing questions.* In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 48043, pp. 307-317).

Aurisicchio, M., Bracewell, R. H., & Wallace, K. M. (2006, January). *Characterising in detail the information requests of engineering designers.* In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 42584, pp. 219-230).

Aurisicchio, M., Bracewell, R., & Wallace, K. (2010). *Understanding how the information requests of aerospace engineering designers influence information-seeking behaviour.* Journal of Engineering Design, 21(6), 707-730.

Batch, A., & Elmqvist, N. (2017). *The interactive visualization gap in initial exploratory data analysis.* IEEE transactions on visualization and computer graphics, 24(1), 278-287.

Baya, V., (1996). *Information Handling Behaviour of Designers during Conceptual Design: Three Experiments, (Doctoral dissertation), Stanford University, Stanford, pp.* 167.

Behrens, J. T. (1997). *Principles and procedures of exploratory data analysis.* Psychological Methods, 2(2), 131.

Berger, W. (2014). *A more beautiful question: The power of inquiry to spark breakthrough ideas.* Bloomsbury Publishing USA.

Bertin, J. (1983). *Semiology of graphics; diagrams networks maps (No.* 04; QA90, B7.).

Bogers, S., Frens, J., Van Kollenburg, J., Deckers, E., & Hummels, C. (2016, June). *Connected baby bottle: A design case study towards a framework for data-enabled design.* In Proceedings of the 2016 ACM Conference on Designing Interactive Systems (pp. 301-311).

Bornakke, T., & Due, B. L. (2018). *Big–Thick Blending: A method for mixing analytical insights from big and thick data sources.* Big Data & Society, 5(1), 2053951718765026.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Using vision to think.* Readings in information visualization: using vision to think, 579-581

Cardoso, C., Eris, O., Badke-Schaub, P., & Aurisicchio, M. (2014). *Question asking in design reviews: how does inquiry facilitate the learning interaction?.*

Cardoso, C., Hurst, A., & Nespoli, O. G. (2020). *Reflective inquiry in design reviews: The role of question-asking during exchanges of peer feedback.* The International journal of engineering education, 36(2), 614-622.

Choi, I. K., Childers, T., Raveendranath, N. K., Mishra, S., Harris, K., & Reda, K. (2019, May). *Concept-driven visual analytics: an exploratory study of model-and hypothesis-based reasoning with visualizations.* In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

Churchill, E. F. (2012). *From data divination to data-aware design interactions, 19(5), 10-13.*

Comment. (n.d.-b). *In Cambridge Dictionary*. Cambridge University Press. Retrieved November 15, 2021, from https://dictionary.cambridge.org/es/diccionario/ingles/comment

Cox, K., Grinter, R. E., Hibino, S. L., Jagadeesan, L. J., & Mantilla, D. (2001). *A multi-modal natural language interface to an information visualization environment*. International Journal of Speech Technology, 4(3), 297-314.

Crilly, N. (2010). *The structure of design revolutions: Kuhnian paradigm shifts in creative problem solving*. Design issues, 26(1), 54-66.

Cross, N. (1997). *Creativity in design: analyzing and modeling the creative leap*. Leonardo, 311-317.

Cross, N. (2004). *Expertise in design: an overview*. Design Studies, 25(5), 427–441.

Cross, Nigel (1982). *Designerly ways of knowing*. Design Studies, 3(4) pp. 221–227.

Data Centric Design Lab. (2020, March 24). *Data is becoming a critical design material*.

Davenport, T. H., Kozyrkov, C., Glaesserr, E., & Siegel, E. (2020). *Strategic Analytics: The Insights You Need from Harvard Business Review*. Harvard Business Press.

Dorst, K. (2011). *The core of design thinking and its application*. Design studies, 32(6), 521-532.

Dorst, K. (2019). *Co-evolution and emergence in design*. Design Studies, 65, 60-77.

Dorst, K., & Cross, N. (2001). *Creativity in the design process: co-evolution of problem solution*. Design studies, 22(5), 425-437.

Dove, G. and Jones, S. (2013). *Evaluating creativity support in co-design workshops*. Paper presented at the CHI 2013 Workshop: Evaluation Methods for Creativity Support Environments, 27th April - 2nd May 2013, Paris, France.

Dove, G. and Jones, S. (2014). *Using Information Visualization to Support Creativity in Service Design Workshops*. Paper presented at the ServDes. 2014, 09-04-2014 - 11-04-2014, Lancaster, UK.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). *Engineering design thinking, teaching, and learning*. Journal of engineering education, 94(1), 103-120.

Elaad. (2021). *Analyses*. Elaad Analyses. https://platform.elaad.io/analyses.html

Elisondo, R. (2016). *Creativity is Always a Social Process*. Creativity. Theories ñ Research -Applications, 3(2), 194-210.

Erickson, B., & Nosanchuk, T. (1992). *Understanding data.* McGraw-Hill Education (UK).

Eris, O. (2004). *Effective inquiry for innovative engineering design (Vol.* 10). Springer Science & Business Media.

Eris, O., (2002). *Perceiving, comprehending, and measuring design activity through the questions asked while designing.* (Doctoral dissertation). Stanford University, pp. 184

Eris, O., Sheppard, S., & Kwan, A. (2007). *Learning to Question: a Longitudinal Analysis of Question Asking Behaviour and Engagement in Project-Based Learning.* In DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design, Paris, France, 28.-31.07. 2007 (pp. 303-304).

Farzaneh, H. H., Kaiser, M. K., & Lindemann, U. (2012). *Creative Processes in GroupsñRelating Communication, Cognitive Processes and Solution Ideas.* In DS 73-1 Proceedings of the 2nd International Conference on Design Creativity Volume 1 (pp. 13-22).

Fayyad, U. M., Wierse, A., & Grinstein, G. G. (2002). *Information Visualization in Data Mining and knowledge discovery.* Morgan Kaufmann.

Fekete, J. D., Van Wijk, J. J., Stasko, J. T., & North, C. (2008). *The value of information visualization.* In Information visualization (pp. 1-18). Springer, Berlin, Heidelberg.

Few, S. (2014). *Data visualization for human perception.* The Encyclopedia of Human-Computer Interaction, 2nd Ed.

Fiaz, A. S., Asha, N., Sumathi, D., & Navaz, A. S. (2016). *Data visualization: enhancing big data more adaptable and valuable.* International Journal of Applied Engineering Research, 11(4), 2801-2804.

Ford Authority. (2017, September 12). *Ford's Ideo Partnership Echoes CEO Jim Hackett's Steelcase Strategy.*

Gao, T., Dontcheva, M., Adar, E., Liu, Z., & Karahalios, K. G. (2015, November). *Datatone: Managing ambiguity in natural language interfaces for data visualization.* In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (pp. 489-500).

Gershon, N., Eick, S. G., & Card, S. (1998). *Information visualization.* interactions, 5(2), 9-15.

Gino, F., Todorova, G., Miron-Spektor, E., & Argote, L. (2009). *When and why prior task experience fosters team creativity.* In Creativity in groups. Emerald Group Publishing Limited

Gorkovenko, K., Burnett, D. J., Thorp, J. K., Richards, D., & Murray-Rust, D. (2020, April). *Exploring the future of data-driven product design.* In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

Gotz, D., & Zhou, M. X. (2009). *Characterizing users' visual analytic activity for insight provenance.* Information Visualization, 8(1), 42-55.

Gotz, D., Zhou, M. X., & Aggarwal, V. (2006, October). *Interactive visual synthesis of analytic knowledge.* In 2006 IEEE Symposium on Visual Analytics Science And Technology (pp. 51-58). IEEE.

Graesser, A., Rus, V., & Cai, Z. (2008). *Question classification schemes.* In Proc. of the Workshop on Question Generation (pp. 10-17).

Grebici, K., Aurisicchio, M., & Bracewell, R. (2009, January). *Guiding engineering design activities through a question based approach.* In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 48999, pp. 1449-1459).

Green, T. M., Ribarsky, W., & Fisher, B. (2009). *Building and applying a human cognition model for visual analytics.* Information visualization, 8(1), 1-13.

Gruber, T. R. and Russel, D. M., 1992. Derivation and Use of Design Rationale Information.

Guerreiro Goncalves, M. (2016). *Decoding designers' inspiration process.* (Doctoral dissertation). Faculty of Industrial Design Engineering. TU Delft

Hanington, B. (2003). *Methods in the making: A perspective on the state of human research in design.* Design issues, 19(4), 9-18.

Hao, S. (2021). *Data as a creative material supporting ideation in Ford: creative problem solving with data.* (Master thesis). Faculty of Industrial Design Engineering. TU Delft

Henderson, K. (1998). *On line and on paper: Visual representations, visual culture, and computer graphics in design engineering.* MIT press

Henschen, D. (2017) *Analytics Team Democratizes Data-Driven Analysis.*

Hui, Q., Li, Y., Tao, Y., & Liu, H. (2020). *Triple-Helix Structured Model Based on Problem-Knowledge-Solution Co-evolution for Innovative Product Design Process.* Chinese Journal of Mechanical Engineering, 33(1), 1-15.

Hui, Q., Li, Y., Tao, Y., & Liu, H. (2020). *Triple-Helix Structured Model Based on Problem-Knowledge-Solution Co-evolution for Innovative Product Design Process.* Chinese Journal of Mechanical Engineering, 33(1), 1-15.

Hurst, A., Duong, C., Flus, M., Litster, G., Nickel, J., & Dai, A. (2021, July). *Evaluating Peer-led Feedback in Asynchronous Design Critiques: A Question-centered Approach*. In 2021 ASEE Virtual Annual Conference Content Access.

IDEO & Ford. (2019, December 4). *How Design is Driving Ford to Reimagine What a Car Company Can Be*. IDEO.com

IDEO. (2020, June 26). *What is Design Thinking? IDEO U.*

Jaarsveld, S., Lachmann, T., Hamel, R., & Leeuwen, C. V. (2010). *Solving and creating raven progressive matrices: reasoning in well-and ill-defined problem spaces.* Creativity Research Journal, 22(3), 304-319.

Jansen, P. (2021). *Implementation strategy of a data enabled service design process at Ford (Masterr thesis)*. Faculty of Industrial Design Engenieering. TU Delft

Jebb, A. T., Parrigon, S., & Woo, S. E. (2017). *Exploratory data analysis as a foundation of inductive research*. Human Resource Management Review, 27(2), 265-276.

Jewell, D., Barros, R. D., Diederichs, S., Duijvestijn, L. M., Hammersley, M., Hazra, A., ... & Zolotow, C. (2014). *Performance and capacity implications for big data*. IBM Redbooks.

Jung, J., Kleinsmann, M., & Snelders, D. (2019, September). *Reviewing Design Movement Towards the Collective Computing Era: How will Future Design Activities Differ from Those in Current and Past Eras of Modern Computing*. In International Association of Societies of Design Research Conference.

Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). *Visual analytics: Definition, process, and challenges*. In Information visualization (pp. 154-175). Springer, Berlin, Heidelberg.

King, R., Churchill, E. F., & Tan, C. (2017). *Designing with data: Improving the user experience with A/B testing*. O'Reilly Media, Inc.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons.

Koslowski, T. (2012) *How technology is ending the automotive industry century-old business model. Gartner Research (September 25, 2012)*

Kotchka, C. (2019, January 16). *Why Creativity Is More Important than Ever.* IDEO U.

Krum, R. (2013). *Cool infographics: Effective communication with data visualization and design*. John Wiley & Sons.

Kuffner, T. A. and Ullman, D. G., (1991). *The Information Request of Mechanical Design Engineers*. Design Studies, Vol. 12, No. 1, pp. 42-50.

Kun, P. (2020). *Design Inquiry Through Data.*

Kun, P., Mulder, I., & Kortuem, G. (2018, July). *Design enquiry through data: appropriating a data science workflow for the design process.* In Proceedings of the 32nd International BCS Human Computer Interaction Conference 32 (pp. 1-12).

Lawson, B. (1979). *Cognitive strategies in architectural design.* Ergonomics, 22(1), 59 - 68.

Lawson, B. (2004). *Schemata, gambits and precedent: some factors in design expertise.* Design Studies, 25(5), 443–457.

Leinhardt, G., & Leinhardt, S. (1980). *Chapter 3: Exploratory data analysis: New tools for the analysis of empirical data.* Review of research in education, 8(1), 85-157.

Lloyd, P., & Scott, P. (1994). *Discovering the design problem.* Design Studies, 15(2), 125–140.

Lubart, T. I. (2001). *Models of the creative process: Past, present and future.* Creativity research journal, 13(3-4), 295-308.

Maher, M. L., Poon, J., & Boulanger, S. (1996). *Formalizing design exploration as co-evolution.* In Advances in formal design methods for CAD (pp. 3-30). Springer, Boston, MA.

Malaga, R. A. (2000). *The effect of stimulus modes and associative distance in individual creativity support systems.* Decision Support Systems, 29(2), 125-141.

Marti, P., Megens, C. and Hummels, C. (2016). *Data-enabled design for social change: Two case studies.* Future Internet, 8(4), 46.

McGinley, C., & Dong, H. (2011). *Designing with information and empathy: Delivering human information to designers.* The Design Journal, 14(2), 187-206.

McKenna, S., Mazur, D., Agutter, J., & Meyer, M. (2014). *Design activity framework for visualization design.* IEEE Transactions on Visualization and Computer Graphics, 20(12), 2191-2200.

McKinsey & Company (2016). *Monetizing car data: new service business opportunities to create new customer benefits.* McKinsey & Company.

McKinsey & Company (2019). *Race 2050: A Vision for the European Automotive Industry.* McKinsey Center for Future Mobility, retrieved on 4th May.

McNeil, D. R., McNeil, D. K., & McNeil, D. R. (1977). *Interactive data analysis: a practical primer.* John Wiley & Sons.

Mednick, S. (1962). *The associative basis of the creative process.* Psychological review, 69(3), 220.

Moore, D. S., & McCabe, G. P. (1989). *Introduction to the Practice of Statistics.* WH Freeman/Times Books/Henry Holt & Co

Moreau, C. P., & Engeset, M. G. (2016). *The downstream consequences of problem-solving mindsets: How playing with LEGO influences creativity*. Journal of Marketing Research, 53(1), 18-30.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: a second course in statistics.*

Muller, W. (1989). *Design discipline and the significance of visuo-spatial thinking*. Design Studies, 10(1), 12-23.

Mumford, M. D., & Antes, A. L. (2007). *Commentaries: Debates About the General Picture: Cognition and Creative Achievement*. Creativity Research Journal, 19(4), 367-374.

Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Nightingale, P. (2009). *Tacit knowledge and engineering design*. In Philosophy of technology and engineering sciences (pp. 351-374). North-Holland.)

North, C. (2006). *Toward measuring visualization insight*. IEEE computer graphics and applications, 26(3), 6-9.

Oxford University Press (OUP). (*n*.d.). Definition of cognition. Lexico.Com. Retrieved September 25, 2021, from https://www.lexico.com/definition/cognition

Peng, R. (2012). *Exploratory data analysis with R*. Lulu.

Petre, M. (2004). *How expert engineering teams use disciplines of innovation*. Design Studies, 25(5), 477-493.

Pike G., Edgar G., Edgar H. (2012). *Perception*. In: Braisby, Nick and Gellatly, Angus eds. Cognitive Psychology. Oxford University Press, pp. 65–99.

Plsek, P. E. (1996). *Working paper: Models for the creative process*. Directed Creativity

PwC Strategy (2020) *The 2020 Digital Auto Report. Navigating through a post-pandemic world*. Volume 1.

Question. (n.d.-b). *In Cambridge Dictionary*. Cambridge University Press. Retrieved November 15, 2021, from https://dictionary.cambridge.org/es/diccionario/ingles/question

Ragan, E. D., Endert, A., Sanyal, J., & Chen, J. (2015). *Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes*. IEEE transactions on visualization and computer graphics, 22(1), 31-40.

Reda, K., Johnson, A. E., Papka, M. E., & Leigh, J. (2016). *Modeling and evaluating user behavior in exploratory visual analysis*. Information Visualization, 15(4), 325-339.

Rhodes, M. (1961). *An analysis of creativity*. The Phi delta kappan, 42(7), 305-310.

Rowley, J. (2007). *The wisdom hierarchy: representations of the DIKW hierarchy*. Journal of information science, 33(2), 163-180.

Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., & Keim, D. A. (2014). *Knowledge generation model for visual analytics*. IEEE transactions on visualization and computer graphics, 20(12), 1604-1613.

Sedig, K., Parsons, P., & Babanski, A. (2012). *Towards a characterization of interactivity in visual analytics*. J. Multim. Process. Technol., 3(1), 12-28.

Setlur, V., Battersby, S. E., Tory, M., Gossweiler, R., & Chang, A. X. (2016, October). *Eviza: A natural language interface for visual analysis*. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (pp. 365-377).

Shrinivasan, Y. B., & van Wijk, J. J. (2008, April). *Supporting the analytical reasoning process in information visualization*. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1237-1246).

Spalburg, T. (2021). *Development of the Ford Concreate, a data immersion toolkit: Building data for better insights*. (Master thesis). Faculty of Industrial Design Engineering. TU Delft

Srinivasan, A., & Stasko, J. (2017, June). *Natural language interfaces for data analysis with visualization: Considering what has and could be asked*. In Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers (pp. 55-59).

Stasko, J., Gorg, C., & Liu, Z. (2008). *Jigsaw: supporting investigative analysis through interactive visualization*. Information visualization, 7(2), 118-132.

Steele, J., & Iliinsky, N. (2010). *Beautiful visualization: looking at data through the eyes of experts*. O'Reilly Media, Inc.

Sun, Y., Leigh, J., Johnson, A., & Lee, S. (2010, June). *Articulate: A semi-automated model for translating natural language queries into meaningful visualizations*. In International Symposium on Smart Graphics (pp. 184-195). Springer, Berlin, Heidelberg.

Tardif, T. Z., & Sternberg, R. J. (1988). *17 What do we know about creativity?*. The nature of creativity: Contemporary psychological perspectives, 429.

Tergan, S. O., & Keller, T. (Eds.). *(2005)*. Knowledge and information visualization: Searching for synergies (Vol. 3426). Springer.

Torrance, E. P. (1988). *The nature of creativity as manifest in its testing*. The nature of creativity, 43-75.

Tufte, E. (2001). *The visual display of quantitative information.*

Tukey, J. W. (1977). *Exploratory data analysis (Vol.* 2, pp. 131-160).

Tukey, J. W. (1980). *We need both exploratory and confirmatory.* The American Statistician, 34(1), 23-25.

Van Boeijen, A., Daalhuizen, J., van der Schoor, R., & Zijlstra, J. (2014). *Delft design guide: Design strategies and methods.*

van Breemen, D. (2019).*The use of data visualization during the creative process.* (Master thesis) Communication and Information Sciences, Specialization Communication Design, Faculty of Humanities and Digital Sciences, Tilburg University

Van Wijk, J. J. (2005, October). *The value of visualization.* In VIS 05. IEEE Visualization, 2005. (pp. 79-86). IEEE.

Wallas, G. (1926). *The art of thought.* J. Cape: London.

Walny, J., Frisson, C., West, M., Kosminsky, D., Knudsen, S., Carpendale, S., & Willett, W. (2019). *Data changes everything: Challenges and opportunities in data visualization design handoff.* IEEE transactions on visualization and computer graphics, 26(1), 12-22.

Ware, C. (2005). *Visual queries: The foundation of visual thinking.* In Knowledge and information visualization (pp. 27-35). Springer, Berlin, Heidelberg.

Ware, C. (2019). *Information visualization: perception for design.* Morgan Kaufmann.

Welter, M. M., Jaarsveld, S., & Lachmann, T. (2017). *Problem space matters: The development of creativity and intelligence in primary school children.* Creativity Research Journal, 29(2), 125-132.

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). *Creating an understanding of data literacy for a data-driven society.* The Journal of Community Informatics, 12(3).

Wongsuphasawat, K., Qu, Z., Moritz, D., Chang, R., Ouk, F., Anand, A., ... & Heer, J. (2017, May). *Voyager 2: Augmenting visual analysis with partial view specifications.* In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 2648-2659).

Yi, J. S., Kang, Y. A., Stasko, J. T., & Jacko, J. A. (2008, April). *Understanding and characterizing insights: how do people gain insights using information visualization?.* In Proceedings of the 2008 Workshop on beyond time and errors: novel evaluation methods for Information Visualization (pp. 1-6).

Ziemkiewicz, C., & Kosara, R. (2007). *Understanding information visualization in the context of visual communication.*