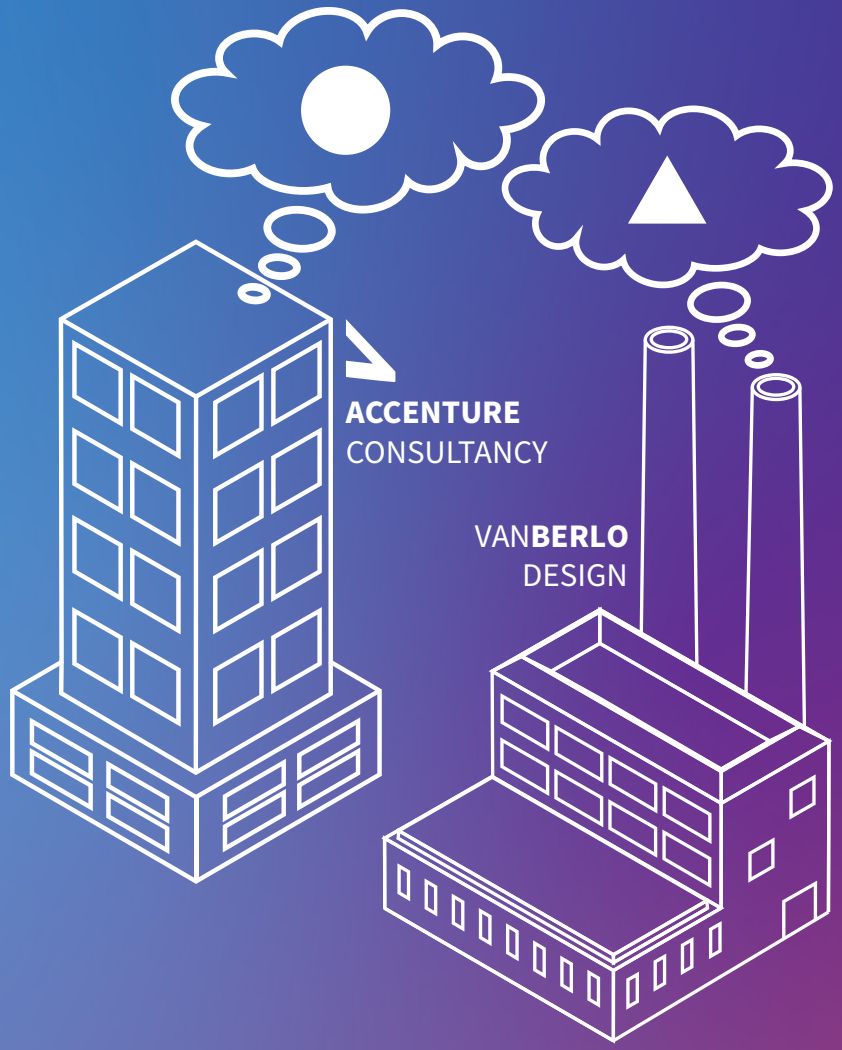


CO-CREATING ORGANIZATIONAL SYNERGY

REDESIGNING A SYNERGY CREATION PROCESS FOR ACCENTURE AND VANBERLO

MATER THESIS BY JORIS RAAPHORST
INDUSTRIAL DESIGN

IN COLLABORATION WITH
TU/E & ACCENTURE NETHERLANDS

2024

CO-CREATING ORGANIZATIONAL SYNERGY

Redesigning A Synergy Creation Process For Accenture And Vanberlo

Author
Joris Raaphorst

Master's Thesis
MSc. Industrial Design
Department of Systemic Change
Technical University of Eindhoven

Assessment Panel
First Examiner & Mentor / Lu, Y.
Expert Examiner / Kuijer, S.C.
Independent Examiner / Essen van, H.A.

Company Supervisor
Leatitia van Wijnen & Rens van Mens

July 2024

ABREVIATIONS

BDM Business Dynamics Modelling
BE Business and Entrepreneurship
CA Creativity and Aesthetics
CC Context Capturing
FI Frame Innovation
H&PS Health and Public Services
ID Industrial Design
IX Industry X
JVP Joint Value Proposition
MDC Math, Data and Computing
OI Opportunity Identification
OM Opportunity Modelling
OS Opportunity Selection
RQ Research Question
S&C Strategy and Consulting
SCP Synergy Creation Process
SS Synergy Strategy
TL Task Load
TLX Task Load Index
TR Technology and Realization
TS Task Specific
UEQ User Experience Questionnaire
US User and Society
UX User Experience
VB VanBerlo

Figure 1 - Impression of use-case introduction presentation at the VanBerlo Eindhoven Office

ABSTRACT

This master project by Joris Raaphorst at the Industrial Design (ID) department TU/e, extends his 2023 study, supporting the acquisition and merger of Design Agency VanBerlo* (Figure 1 & 2) into the IT consultancy Accenture. This study was conducted from within Accenture - Industry X (IX).

The initial study aimed to create a process to support mutual understanding to drive the operational and cultural synergy between VanBerlo and Accenture Industry X, resulting in a three-phase 'synergy creation process' (SCP) that (1) collected contextual indicators from various perspectives, and (2) synthesizes them into a visual mapping to (3) drive an explorative, co-creative synergy modeling workshop. In this process, mainly based on a too-high participant workload, various areas of improvement were identified.

This study started with a re-design of the three-phase process with a focus on reducing task load (TL) and improving user experience (UX) following a literature-inspired approach resulting in a five-phase design. From these phases, three were updated by replacing or integrating validated techniques.

This was then tested in a pilot based on which the overall layout and explainers were streamlined and the difficulty, mental, and temporal demands of the tasks per session were balanced. Then, the design was evaluated in the context of a real business case on TL and six UX indicators: (1) Attractiveness, (2) Perspicuity, (3) Efficiency, (4) Dependability, (5) Stimulation, and (6) Novelty. The reason that effectiveness was not included in the evaluation was that a poor UX and high TL were expected to have a too large impact on the session's effectiveness.

The re-design did significantly reduce TL for each session. The UX evaluation, however, revealed the need for a pre-session communication protocol to improve clarity and to meet the expectations of participants. From the integrated, validated methods, the Qualitative Group Modelling method showed a promising performance. The Modified Delphi method did not meet the expected TL and UX improvements, but participant feedback suggests a positive impact on content quality. The expanded selection methods reported issues about selection criteria that were too broad and flow inefficiencies, suggesting the need for a revision of this phase.

The SCP provides us with a fairly viable (based on TL) and user-friendly approach to explore organizational synergy in a constructive way. However, no claims about the effectiveness of SCP design could be made, some UX metrics (Stimulation sub-scale) and participant feedback suggest that this process was experienced as valuable and interesting.

*VanBerlo was rebranded as Industrial Design in 2023 but is referred to as VanBerlo (VB) to avoid confusion with the university department.

Figure 2 - Impression VanBerlo Eindhoven Office

CONTENTS

INTRODUCTION

Covers the situational context and the previous works this study is building upon

SITUATIONAL CONTEXT

Past decades have been characterized by accelerating economic and technological paradigm shifts (Brand & Rocchi, 2011), companies experienced increased environmental instability, making exploitative actions less effective (Walrave, Oorschot, & Romme, 2010; Jansen, Van Den Bosch, & Volberda, 2006; Uotila, Maula, Keil, & Zahra, 2009). Their success now depends largely on their capability to balance explorative and exploitative activities to adapt to instability (Hannan & Freeman, 1984; Porac J. a., 1990; Tripsas & Gavetti, 2000; Tushman, Newman, & Romanelli, 2004) (Hannan & Freeman, 1984; Porac J. a., 1990; Tripsas & Gavetti, Capabilities, cognition, and inertia: Evidence from digital imaging, 2000; Tushman, Newman, & Romanelli, 2004). And since today's environment is becoming increasingly unstable (e.g., rapid development of disruptive technologies, the sustainability paradigm shift, increased conflicts, a recent global pandemic etc.), the focus is shifting towards exploration (De Matteis, 2016).

This rising need for adaptability has led the market to seek partners offering end-to-end business, technology, and creativity services (Treichler, 2019). This might be the underlying reason for an emerging trend in which, since 2016, large consultancies have increasingly acquired creative and design capabilities (Schultz, 2019; Quinlan, 2018; Bos & Lundberg, 2019; Gianatasio, 2017). Acquisitions are notoriously difficult, of which 70-90% fails (Fantaguzzi & Handscomb, 2024). Reports on these creative acquisitions, however suggest that they are even more challenging (Weber, 2019). The author experienced such an acquisition firsthand, working as a design engineer at the Dutch design agency VanBerlo, which was acquired by Accenture in 2020 (Accenture Newsroom, 2020). And now, four years later, many of the merger challenges remain, and the influx of projects has decreased, while for collaboration, only a few examples can be found.

PREVIOUS WORK

Prior to this study, various other efforts were made to support this merger. A task force was established to market VanBerlo's capabilities internally and rebrand it as Industrial Design. As part of these efforts, a study was conducted by Reus (2023) whose master thesis explored "How the mutual lack of understanding between Industry X's and VanBerlo's capabilities can be bridged to facilitate the creation of synergy during business development?". The resulting Joint Value Proposition (JVP) workshop from Reus (2023) was internally well received, but neither the designed process nor the results were internally adopted. Therefore a follow-up study (M2.1) was conducted, which redesigned the process to improve result adoption. This paper builds upon that study.

Prior Design Objective

The prior study assessed, based on the double-loop learning diagram (Argyris, 1985) that the original JVP workshop primarily focused on capturing information to support opportunity identification. However, the optimized model for learning by Sterman (2000) suggests that the reconfiguration of mental models and seizing opportunities by constructive decision-making is also required to effectively adopt processes and results. This insight is based on the concepts of dynamic capabilities (Teece, Pisano, & Shuen, 1997) and virtual worlds (Schön, 1996; Beckhard & Harris, 1987; Dyer & Dyer, 1994; Michael, 1997; Schein E. H., 1987; Schein E. H., 1988).

Prior Design Framework

Raaphorst (2023) redesigned this process using inspiration from two methodologies. Firstly, to support the reframing of differentiating perspectives, the nine-step process from Frame Innovation (FI) theory (Dorst, 2015) was incorporated.

Secondly, a five-step Business Dynamics Modelling (BDM) approach (Sterman, 2000) was implemented to support 'double-loop learning', by simulating reality in a virtual setting. This approach creates a simplified version of reality, supporting the processing of information feedback loops, providing a basis for decision-making (Porac J. a., 1990; Tripsas & Gavetti, Capabilities, cognition, and inertia: Evidence from digital imaging, 2000; Winter, 2000). It intends to capture shared understanding of a complex challenge in a visual artifact, providing a reference point to create a forum for discussion (Jones, 2017; Scott, 2016; Videira, Antunes, Santos, & Lopes, 2010; 2017). It is seen as an effective and proven tool to tackle dynamic and non-linear challenges (Davis, Eisenhardt, & Bingham, 2007; Romme A. G., 2004; Romme A. G., 2010; Sterman, 2000), making it suitable for the dissemination and continued evolution of identified synergy opportunities. Combining these two frameworks with the Joint Value Proposition workshop from Reus (2023) resulted in the design of a three-phase workshop, 'Synergy Creation Process' (SCP) (Figure 3). This process was tested and refined during three iterative cycles with VanBerlo and Accenture employees.

Initial Workshop Design

Initial Context Capturing Format

The initial session involved a structured one-on-one session with employees from both organizations in which they captured the nature of each group using a business model canvas (BCM) (Osterwalder A. , 2008) and a modified customer journey mapping (CJM) (XO-Projects, 2024) (Appendix 1.1.1) .

Initial Blueprint Creation Process

Performed by the researcher, the collected information was tagged, analyzed thematically, and then organized in a causal and chronological structure to reflect the project/business develop-

PREVIOUS WORK

ment cycle and end-to-end product cycle. The blueprint aimed to highlight similarities and differences in the service offerings of both organizations from the interviews (Appendix 1.2.1).

Initial Synergy Modelling Workshop

Then, involving all earlier participants, a two-hour workshop was conducted, which consisted of ten steps that revolved around the service blueprint (Appendix 1.0).

1. Determine Opportunity & Challenge areas.
2. Formulate Opportunities & Challenges.
3. Select 12 main opportunities for improvement.
4. Discuss causes & Effects.
5. Discuss underlying factors in keywords.
6. Mapping these on the blueprint.
7. Build the story, identify feedback structures.
8. Design & Evaluate Synergy Strategies.
9. Convert to goals.
10. Set actions, ownership & deadlines.

Prior Conclusion

Various parts, especially the modelling tasks, were challenging for the participants, sometimes not able to perform tasks properly. Raaphorst (2023) concluded, based on observations, participant feedback and surveying the task load (TL), that both the interview and the conceptual workshop process were too ambitious, attempting to achieve too much in too little time and reflecting poor user experience.

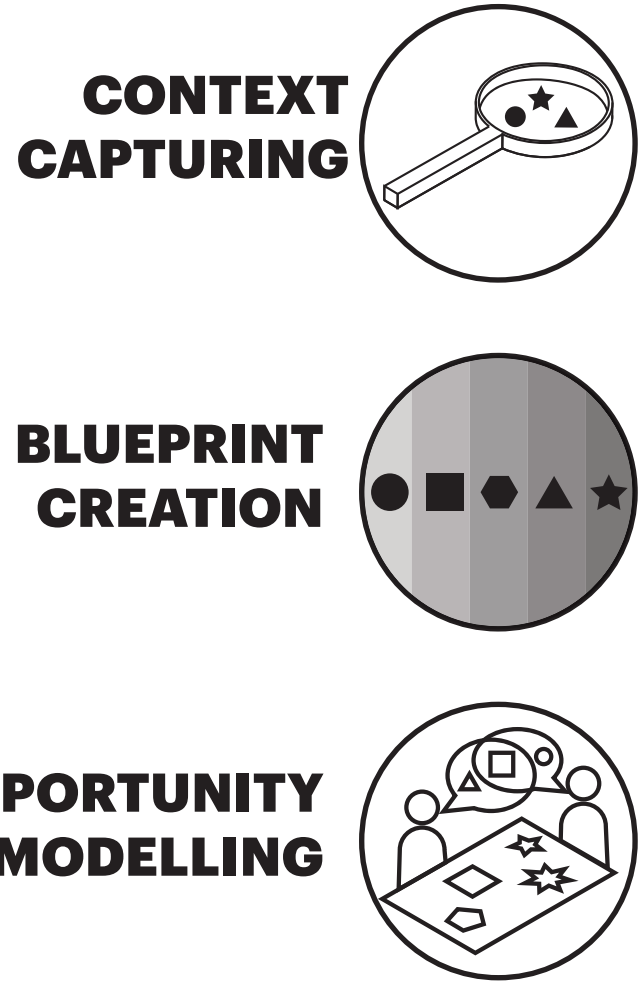


Figure 3 - Initial three-phase process

SCOPE

Explains the research objective and research questions

RESEARCH OBJECTIVES

The antecedent study indicated poor TL and UX on the SCP. Research on process and technology adoption has shown that UX has a significant impact on the attitude toward process and technology adoption (Bruner II & Kumar, 2005; Kapoor, 2014; Kautz, 2004) and that poor UX can lead to resistance and potential rejection (AgeLab, 2020). Therefore, based on Accenture and VanBerlo's need to find synergy in their merger and acquisition process, the primary focus of this study is to research how to balance workload and use UX to increase the adoption of the SCP process and results.

In this study, the SCP is firstly redesigned and then evaluated in a real business context to understand 'how the redesign impacts the task load (TL) and user experience (UX) of the Synergy Creation Process.' To answer this question, this study addressed the TL and UX through two different research questions. The first utilized the data from the previous study to evaluate the effect of the redesign on the participant TL.

1. *"How does the redesign of the SCP impact the task load of each individual session compared to the initial workshop design?"*

The second question compared the UX of the SCP to a general UX benchmark, giving an indication of its overall UX performance.

2. "What is the user experience of participants per session in comparison to the standardized UEQ benchmark?"

Initially, this study also intended to evaluate the effectiveness of the process and results. However, it was found to be challenging to measure the effectiveness of any eventual use behaviors or behavioral intentions without a longitudinal study or evaluating UX-affected attitudinal factors (Venkatesh, Morris, Davis, & Davis, 2003).

To get some indication of effectiveness, the evaluation measured the sessions outputs perceived usability. However, due to an error in data collection, this data was unusable and is further excluded from this study.

METHODOLOGY

Covers the research approach and each of the used research methods

APPROACH

Process
First, problem areas in the SCP design were identified based on the antecedent study. Then, a literature-inspired design process was conducted. This design (Appendix 1) and research protocol (Appendix 2) were first tested in a pilot (Appendix 4) before being tested in a business case study (Appendix 5). The process allowed time to review and revise each piloted session before deployment in the use case (Figure 4).

Problem Identification
The antecedent study findings identified challenges related to effort, process performance, and user experience. Based on these findings, the three-phased SCP was split into five phases; Context Capture (CC), Blueprint Creation (BC), Opportunity Identification (OI), Opportunity Selection (OS), and Opportunity Modelling (OM), providing more time per session.

Literature Inspired Design
For each challenge area, an integrative (Snyder, 2019), pragmatic problem-solution focused literature review, was conducted to identify validated methods and techniques and provide theoretical justification for integration. This entailed flexibly using keywords, relying on systematic reviews from others, and following citations to efficiently find reasonable solutions.

Pilot testing
Pilot studies are viewed as "a crucial element of a good study design" (Van Teijlingen & Hundley, 2002; Hassan, Schattner, & Mazza, 2006), increasing the likelihood of the main study's success.
Van Teijlingen & Hundley (2002) suggest piloting with a low-impact group. Therefore, I started with a group of interns and their supervisors before running the business use case study with high-level internal stakeholders to pilot test both the process design and the research instrument. During the pilot, the complete study run was complemented with open qualitative post-session interview questions to review the survey design, such as: How did you experience the survey? Please point out difficult-to-understand parts. Was anything exhausting? What did you think about the survey length?)

Use-case study
A case study makes the potential output more meaningful and impactful (Feagin, Orum, & Sjoberg, 2016). Testing the design in the use case of the VanBerlo/Accenture merger places it in a situational context of conflicting siloes with high theoretical potential for valuable synergy.

Case Study Selection
Potential case studies were identified with the Senior Managers operating between Industry X and VanBerlo. Based on relevance and participant availability, a pilot and business case study were selected in alignment with Accenture Management.

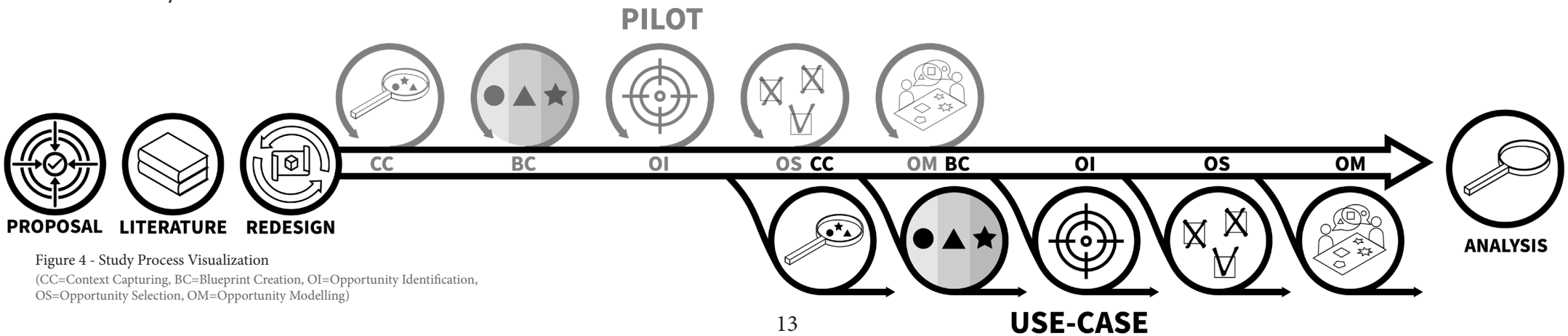


Figure 4 - Study Process Visualization
(CC=Context Capturing, BC=Blueprint Creation, OI=Opportunity Identification, OS=Opportunity Selection, OM=Opportunity Modelling)

DATA COLLECTION

Various data collection methods, such as interviews, focus groups, and observational studies, were considered. These would provide richer qualitative insights (Lune, 2017) but were estimated to be too time-consuming and unfitting for the intended scale (Feagin, Orum, & Sjoberg, 2016, pp. 147-155).

In addition, due to this workshop's potential business and emotional sensitivity, group interview sessions and video/audio recordings were not considered an option to support observations or interviews.

Since the researcher facilitated the workshop, it would have been impossible to reliably record observations. Therefore, to achieve structured, scalable, and granular data collection, the four co-creative sessions of the pilot and use case were evaluated using a combination of post-task and post-session standardized questionnaires (Appendix 2).

Post-Task Effort Evaluation
The perceived task effort was evaluated using a repeated, post-task inquiry of Difficulty (Single Ease Question (SEQ)) (Tedesco & Tullis, 2006; Sauro J. &., 2009), Mental & Temporal Demand (NASA Task Load Index (Hart, 1986)), and an open question prompting participants to explain any low scores below 4 (out of 10).

Using this short post-task inquiry resulted in granular, (Butler, Karpicke, & Roediger III, 2007) quantitative data enriched with qualitative insights about the perceived task effort. This Task Specific (TS) inquiry was conducted three to five times per session, to test each distinct task participants performed.

Post-Session Evaluation
After a completed session, eight metrics were evaluated by combining (1) the five-item ten-point scale NASA Task Load Index (TLX) (Hart, 1986)) and the 32-item seven-point scale, User Experience Questionnaire (UEQ) (Laugwitz, Held, & Schrepp, 2008) measuring UX via seven metrics; (2) Attractiveness, (3) Perspicuity, (4) Efficiency, (5) Dependability, (6) Stimulation, (7) Novelty, and (8) Usability (later excluded due to unreliable data)

Nasa Task Load Index (TLX)
The TLX is a human-centered, validated, widely used (over 82.900 citations), standardized survey for self-reporting workload and perceived performance. Developed initially to evaluate Human-Computer interactions in critical situations, it has, 20 years later, seen a wide range of applications, such as in decision-making, teamwork, and communications (Hart, 2006, October).

The antecedent study used workload data because it was estimated to be a significant challenge. Since this data is available and the workload challenge remains unresolved, it can be used as a benchmark, allowing for a direct comparison to the original design.

User Experience Questionnaire (UEQ)
For the User Experience, various methods were considered, such as the System Usability Scale (Brooke, 1996), Net Promotor Score (Reichheld, 2011), Task Success Rate (Sauro & Lewis, 2016), and Post-Study System Usability Questionnaire (Lewis, 1995).
Eventually, the User Experience Questionnaire (UEQ) (Laugwitz, Held, & Schrepp, 2008) was selected. The UEQ is one of the most recognized questionnaires for UX (Lallemand & Koenig, 2017; Baumgartner & Sonderegger, 2019; Forster, Hergeth, Naujoks, & Krems, 2018; Klammer & van den Anker, 2018), and according to

DATA COLLECTION

a meta-study, since 2017, in Europe, the most commonly used UX questionnaire (Díaz-Oreiro, López, Quesada, & Guerrero, 2019). This study also found that about 10% of the reviewed studies in the meta-analysis that used the UEQ also used the NASA TLX questionnaire.

The UEQ is designed to be a holistic measure of UX (Hassenzahl, 2001) and assesses three pragmatic qualities (Perspicuity, Efficiency, and Dependability), two hedonic qualities (Stimulation and Originality), and Attractiveness. For these criteria, it provides a benchmark based on a "data set containing data from 21175 persons from 468 studies concerning different products (business software, web pages, web shops, social networks)" (Laugwitz, Held, & Schrepp, 2008).

Together, the wide support for this method, the availability of the benchmark data, and the holistic nature of the qualities were the reasons the UEQ was selected to evaluate the SCP design.

Usability (UEQ+)
During the creation and continued development of the UEQ, many additional qualities were developed and validated, which are captured in the UEQ+. One such quality is usability, which was adopted to complement the TLX and UEQ. However, no benchmark for this data was available, and due to an implementation error, the data was unreliable and was not considered in the analysis.

Qualitative Written Feedback
To improve qualitative understanding, three open, written, end-of-survey questions (tips, tops, and other remarks) were included to capture overall feedback (Figure 5).

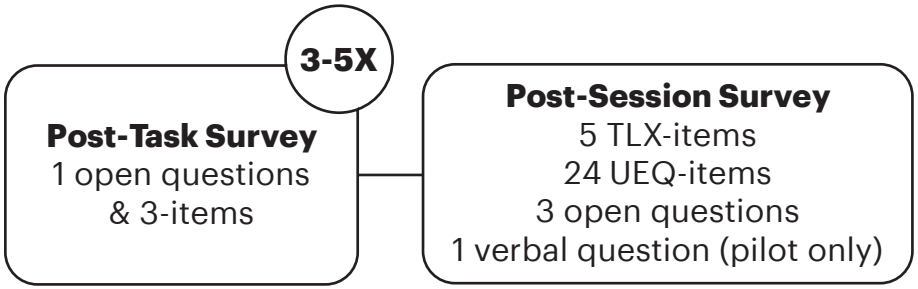


Figure 5 - Data Collection Protocol

STUDY POPULATION

As the SCP addresses change management challenges, it is essential to involve different departmental perspectives from different seniority levels (Cummings & Worley, 2016). In addition, Sterman stressed the involvement of key decision-makers during the process to secure the adoption of system changes, and domain experts to include optimal know-how when modeling system dynamics (Sterman, 2000; Vennix, Andersen, & Richardson, 1997; Lane, 2000). Thus, to improve the quality of the workshop's output, the study purposely takes a maximum variety approach (Hassan M. , 2023) to seniority. Additionally, expert sampling was applied to the OS and OM session, involving key decision-makers (< level 8) during the selection process and opportunity-specific domain experts for the system dynamics modeling.

Pilot Population

The Pilot study consisted of six Industry X group interns and their corresponding six supervisors. Both interns and their supervisors were included to resemble the differences in seniority level. An attempted census approach was taken for the CC, OI, and OS sessions, inviting the complete population of interns and supervisors to participate. A nomination sampling approach during the OS session was used for the OM session to identify opportunity-specific experts for the modeling session.

Use Case Population

For the use case, we consider the complete population of Accenture Netherlands, which is estimated at around 4.000 (Accenture, 2024). Based on expert interviews conducted by Kageler (2024) a population proportion of 44 professionals operating in the life science domain was identified. These professionals represent a wide range of seniority and originate from five different Accenture departments: (1) Health & Public Services (H&PS), (2) Industry-X (IX), (3) VanBerlo (VB), (4) ESP & (5) Strategy & Consulting (S&C) (Table 1).

Sample Size Calculation

A Confidence Level of 95% and a 5% Margin of Error were used to calculate the required sample size. Then, for the CC, OI, and OM sessions, using the group of 44 participants, being a Population Proportion of 1,1%, resulted in a calculated minimum sample size of 13 (Sample-Size-Calculator, z.d.).

Targeting decision-makers for the OS session only senior participants (< level 8) were selected as population proportion. encompassing a group of 21 participants, leading to a calculated minimum sample size of 10 (Sample-Size-Calculator, z.d.).

INVITED PARTICIPANTS	H&PS	IX	ID	ESP	S&C	Total
CC	4	5	4	4	3	20
OI	6	11	5	14	8	44
OS	3	3	3	3	3	15
OM	3	2	2	4	2	13

Table 1 - Total Population Proportion invited to participate

PARTICIPANT SAMPLING

Sample recruitment

For the CC session interviews, senior representatives (< level 7) from each department were asked to identify the four most knowledgeable representatives from each available seniority level, including Analysts (Level 11), Senior Analysts (Level 10), Consultants (Level 9), Associate Managers (Level 8), Managers (Level 7), Senior Managers (Level 6), and Managing Directors (Level 5). For the OI session, an internal business development event was organized in collaboration with another intern (E. Kageler) to attain access to a large study population. Here, census sampling was attempted, inviting the population proportion of 44 participants. For the OS session, three decision-makers from each department were invited based on estimated relevance to the synergy case. For the OM session, a nomination sampling approach was applied to identify opportunity-specific experts, which resulted in a list of thirteen invited participants (Table 2).

POPULATION SENIORITY LEVEL DISTRIBUTION						
SENIORITY LEVEL	H&PS	IX	ID	ESP	S&C	Total
11	1	1	0	2	3	7
10	0	0	0	1	0	1
9	3	2	1	2	1	9
8	0	0	0	6	0	6
7	1	1	3	2	2	9
6	1	2	1	1	1	6
5	0	3	0	0	1	4
4	0	2	0	0	0	2

Table 2- Total Invited Population Proportion - Seniority Level Distribution

ANALYSIS APPROACH

How were the use cases and participants sourced and how data was collected and analyzed

ANALYSIS APPROACH

The TS, TLX, and UEQ required different data processing and interpretation approaches. For the TLX, the antecedent study could be benchmarked. For the TS and UEQ, no benchmark data linked to the SCP exists. The TS, is therefore, only internally comparing the tasks per session. The UEQ method provides a native, generic benchmark for comparison. These comparisons are made for the pilot to support design refinements. The same benchmarks are used to compare against the use-case results, as the pilot is not fully representative due to contextual differences.

Data Normality and Filtering

To determine normal distribution, the data was analyzed using histograms (Appendix 3.1) and Q-Q plots (Appendix 3.3). For the TS data Paired Point Plots (Appendix 4.x.3 & 5.x.3) were used to visually review the behavior of the data and filter the data for any straight.

TLX – Task Load Index

The TLX benchmark was measured for the Context Capturing (CC) sessions and the Modelling Workshop during the antecedent study (Table 3). These were used to compare against the pilot and use-case data. For the CC-session the data from the benchmark, pilot and TLX were compared. Since the workshop was split into three-sessions, its benchmark was compared against the three different resulting sessions from the pilot and use case.

Benchmark analysis

The benchmark shows that the SCP in the antecedent scored high on all metrics except for the frustration indicator. This metric scored extremely low on the 10-point scale, and therefore requires a careful approach. To account for this extreme, besides the P-value and Effect Size, also the relative position of the frustration score on the scale is considered during analysis.

TASK LOAD BENCHMARK	Mental Demand	Temporal Demand	Performance	Effort	Frustration
Context Capturing	8	6,25	7	7,5	1,25
Modelling Workshop	7	6,25	7,25	6,75	1

Table 3 - Task Load Benchmark

TS – Task Specific

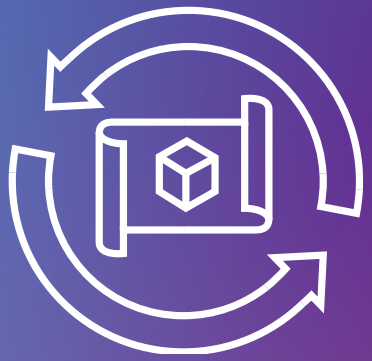
The individual session tasks were only compared internally (per session) on difficulty, mental, and temporal demand to get a better and more granular understanding of how effort distribution is balanced.

TS & TLX data processing

Both the TS & TLX rely on a graded ordinal numerical score that allows for individual or cohort study and statistical analysis (Bell, et al., 2022). Based on the visualized data, using histograms (Appendix 3.2) and Q-Q plots (Appendix 3.3), it was concluded that the assumptions for the interval or normal distribution and homogeneity of variance were not being met, meaning that a non-parametric test is required for data analysis. For the P-values of the TLX the Mann-Whitney U-test was used, as this data covers independent populations (Appendix 4.x.2 & 5.x.2). For the P-values of the internal TS comparison, the Wilcoxon signed-rank-sum test was used since this data is paired (Appendix 4.x.3 & 5.x.3). To support data interpretation the Mean Differences and Effect Sizes were calculated. All TS and TLX data were processed using R-Studio; check Appendix 3.1 for the code used for data processing.

UEQ – User Experience

The UEQ was compared to its native benchmark, (Schrepp, 2017). The UEQ method provides an Excel-based data analysis tool with which a comparison graph is generated, it plots the recorded data on a scale distribution that indicates how your test performed in comparison to the benchmark. The scale distribution distinguishes between five levels: (1) Bad <25%, (2) Below Average <50%, (3) Above Average <75%, (4) Good <90%, and (5) Excellent >90%. The UEQ combines four or six items to provide a more reliable indicator per metric. To assess the data's reliability, the variance per sub-scale is calculated. And to accommodate for small sample sizes, a data-distribution visualization was created per sub-scale to review the behavior of each individual item, providing more insight into the subscale's behavior (Appendix 4.x.4 & 5.x.4).



DESIGN CHANGES

Describes how the design was re-designed and why the changes were necessary

STRUCTURE CHANGES

The high recorded TL of the original three-phase SCP resulted in splitting the original 'Synergy Modelling Workshop' into (initially) four separate sessions of approximately an hour duration, (1) Opportunity Identification (OI), (2) Opportunity Selection (OS), (3) Opportunity Modelling (OM) and (4) Synergy Strategy (SS). Due to the versatility of the adopted Group Modelling method, the modelling and strategy steps were later combined into one two-hour session, resulting in a five-phase SCP Design (Figure 6).

It was hypothesized that this split would reduce the TL per task while remaining viable as it keeps the sessions within feasible time frames that are deployable in a multi-stakeholder corporate context. Next to this, it would allow more strategic incorporation of participants (e.g. juniors, seniors, and decision-makers can be more efficiently distributed across the process), more efficiently utilizing the companies' human resources, improving its viability and potential effectiveness.

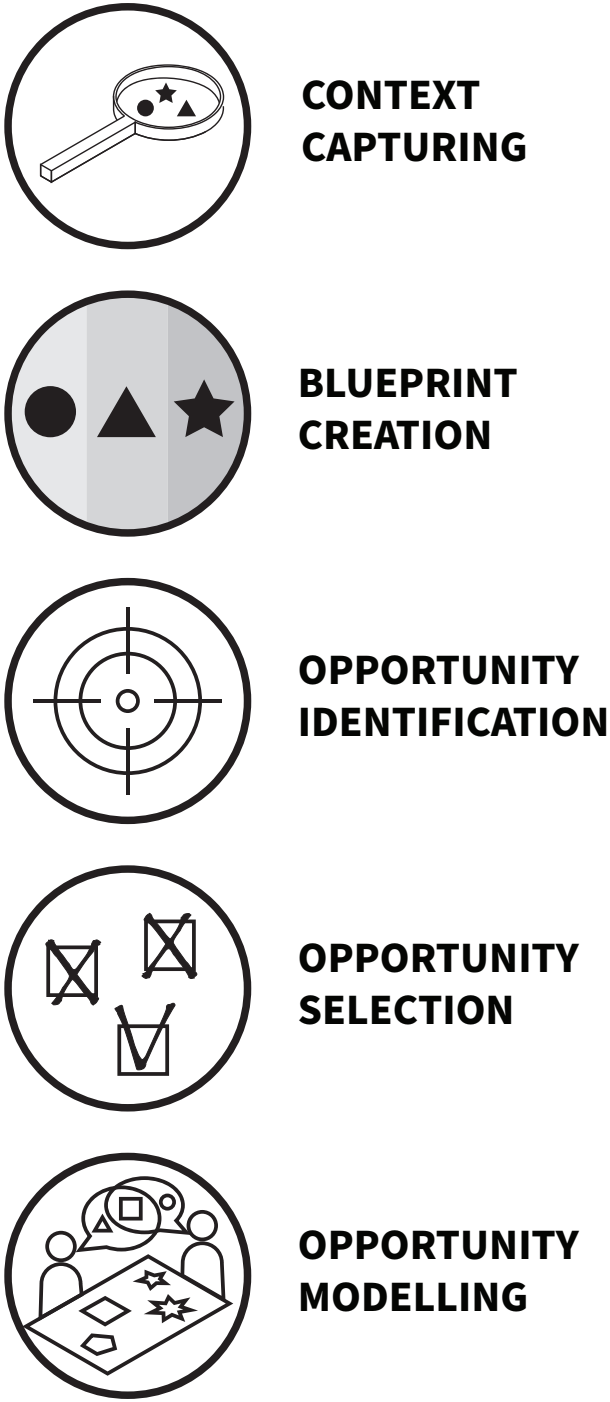


Figure 6 - Five phase SCP process

CONTEXT CAPTURING

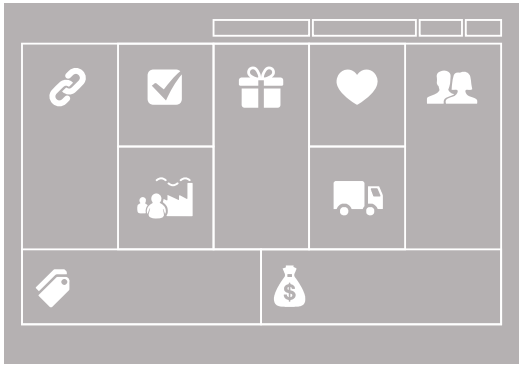
During the antecedent study, it was observed that senior participants working on empty templates spent much time on basic entries that junior participants also included while having more additional insights. Seemingly resulting in higher experienced workloads.

A wide range of Delphi methods (Linstone & Turoff, 2002) were considered. Among them was an approach designed for the context of organizational mergers (Boucher, 1980). It introduced the element of participants reviewing each other's work, a concept Gustafson (1973) validated.

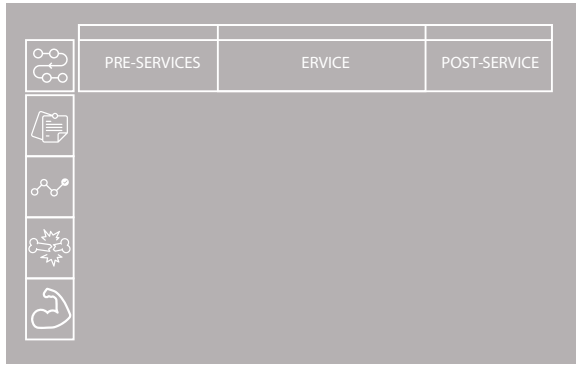
It was hypothesized that using only one data collection template (Appendix 1.1.3) per group (experts reviewing each other's input) would improve the value of session outputs while reducing the workload of the participants for all the reviewing (not first) participants. For the Starting Participants, no difference was expected as the design of their session was comparable to the Raaphorst (2023) benchmark. Additionally, this design change was expected to simplify the blueprint creation process, as it would prevent duplicates, reducing synthesizing efforts.

Lastly, the task to 'capture existing and short-term emerging synergies' was included to provide synergy examples that could be exchanged during the OI-session to mitigate the observed risk of getting distracted preventing the identification of new potential synergy opportunities (Appendix 1.1.2 & 1.1.3) (Figure 7).

1. BUSINESS MODEL CANVAS



2. CUSTOMER JOURNEY MAPPING



3. CAPTURE EXISTING SYNERGIES

Figure 7 - CC session and task design

BLUEPRINT CREATION

As high-volume insights were collected for not two but five perspectives, the original blueprint template no longer sufficed.
In short, the researcher's blueprint creation process was expanded into a five-step protocol (Figure 8). Firstly, a (1) word count analysis was included to support formulating codes for (2) organizing the information in themes. Comparable to the original design, the themes were then (3) mapped on the service cycle in chronological order before the (previously tagged) items were (4) sorted on their dedicated swim lane. Lastly, the content on the template was (5) synthesized by merging similarities into a segment description, leaving only the organizational differences visible on the swim lanes.

Competence integration
The master's program requires the integration of five expertise areas: (1) User and Society, (2) Business and Entrepreneurship, (3) Creativity and Aesthetics, (4) Math data and computing, and (5) Technology and Realization. The blueprint creation process is an example of converging expertise integration and consideration.
The blueprint, inspired by System-Oriented-Design (SOD) (Sevaldson, 2013) and Giga-mapping (Sevaldson, 2018), places the idea of artifact-based ideation (Zimmerman, Forlizzi, & Evenson, 2007) in the complex context, organizational siloes (Tett, 2015), mergers and exploration-exploitation tension (March, 1991; Gupta, Smith, & Shalley, 2006).

It integrates participant-derived expertise captured in the CC session via the BMC & CJM. And required the researcher to combine different departmental, client-facing value propositions (Osterwalder, Pigneur, Bernarda, & Smith, 2015), business development (Fischer, Gebauer, Gregory, Ren, & Fleisch, 2010) and project management processes (Patrício, Costa, Pereira, & António, 2022) into one visual service blueprint.

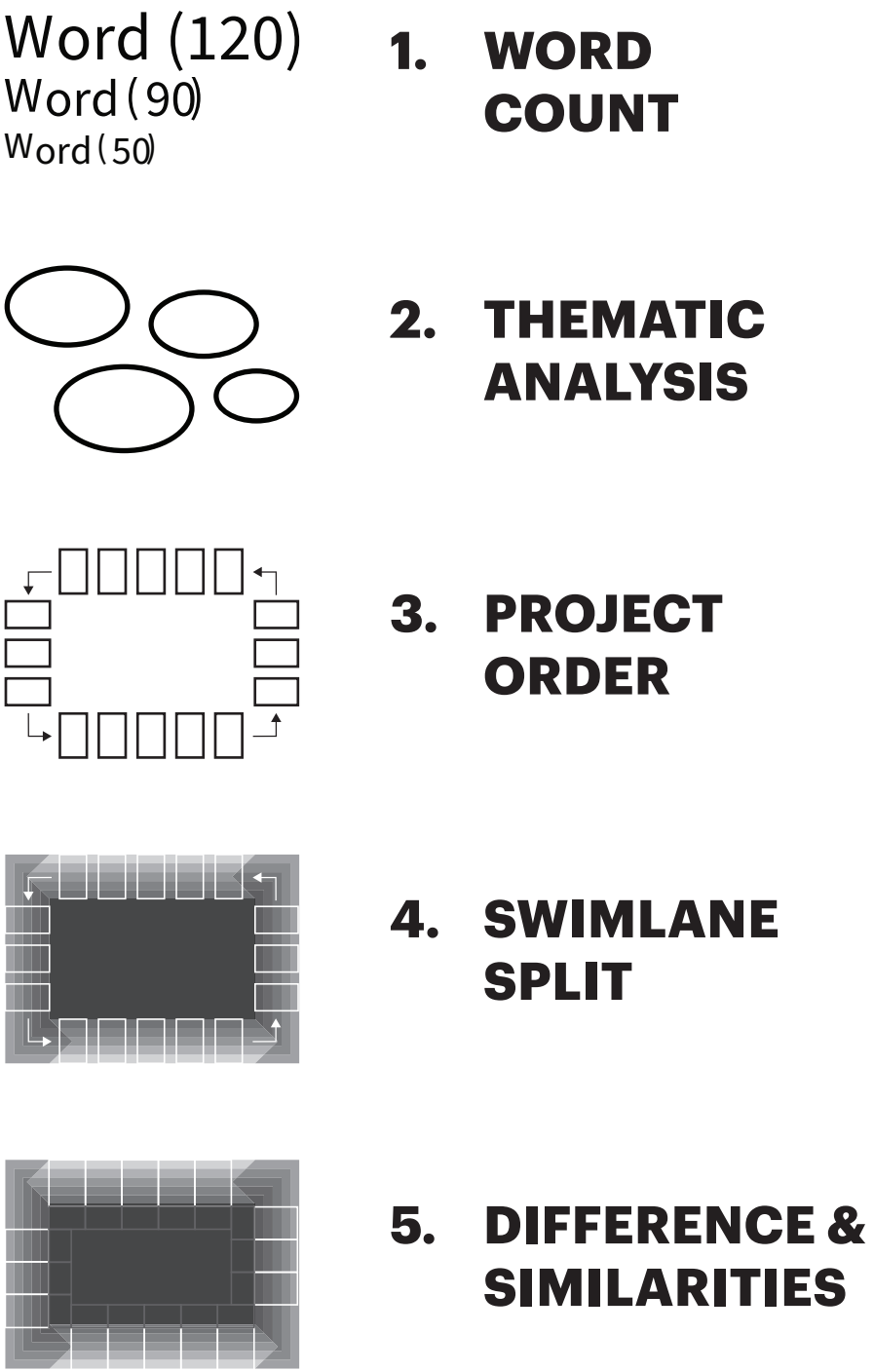


Figure 8 - Blueprint creation process

BLUEPRINT CREATION

Two design changes were made to properly organize and represent the collected information. First, the cyclical 'pre-service, service, post-service' structure inspired by the journey mapping method (Stickdorn, M. E., & Schneider, 2018) was adopted to accommodate all BMC and CJM data while introducing an interesting link between post- and pre-service. Second, to allow the integration of more departments and safeguard aesthetics and understandability, the original causal diagram (Appendix 1.2.1) was replaced with a swim lane diagram (Appendix 1.2.3), which is specialized for operating between different departments and competitors (Rummler & Brache, 2012).

Together, they formed the basis for the blueprint and accommodated most departmental similarities and differences while likely improving attractiveness and cognitive ergonomics compared to the original design (e.g., readability, understandability & mental demand).

To populate the blueprint, a thematic analysis (Braun & Clarke, 2012) inspired approach was taken. The creation of the codebook was supported by a computer-based word frequency analysis (Stone, 2020) to make the high information volume manageable. This resulted in 109 groups (Confidential Appendix 7.2.1 & 7.2.2), which were then placed on the blueprint and organized in chronological project order (Confidential Appendix 7.2.3). After this, the groups were combined into phases (themes) on the template (Confidential Appendix 7.2.4).

Further, driven by participant feedback from the prior study, the information volume was reduced to improve readability by applying additional synthesis layers to the data conversion. Per group, the content was compared between departments, and similarities were extracted and added to a phase description, thus only highlighting the differences between the departments (Confidential Appendix 7.2.5).

In addition, this anonymized the information guarding the privacy and safety of the participants. This was further supported by a facilitator and workshop protocol with guidelines to safeguard and respect socio-cultural differences (Appendix 1.0.1).

Lastly, due to the geographic distance between participants, a primarily two-dimensional format was used, allowing both online and physical hosting. Physical sessions were, however, preferred for socialization reasons (Bauer & Green, 1994). Therefore, physical tabletop tools were created to support facilitation.

To resemble the versatility of online environments like Miro (Miro, sd), a transparent polycarbonate (PC) sheet was placed over the blueprint, allowing marker annotations. In addition, thin, transparent polyvinylchloride (PVC) hexagons and rectangles were added to facilitate the movability of ideas and information while preventing the template from becoming invisible (Appendix 1.3.6).

OPPORTUNITY IDENTIFICATION

The objective of the first part of the split synergy workshop was a one-hour session that generated a large volume of properly defined opportunities and incorporated the 'Existing and Emerging Synergy Efforts' in a way that stimulates the participants to expand awareness about existing synergies, allowing them to focus on exploring new synergies.

The redesigned OI-process consisted of five main tasks aimed at fostering collaboration and innovation (Figure 9). As per the original design, participants began by (1) identifying areas of interest (Appendix 1.3.1), allowing them to familiarize themselves with the blueprint. Then, a new task was integrated of (2) positioning existing and emerging synergies (Appendix 1.3.2) to support further familiarization with the blueprint and provide examples of successful synergy cases. This step helps to mitigate repetition and inspires more grounded ideas. The third task remained unchanged, which was (3) identifying new opportunities (Appendix 1.3.3), this was first conducted individually and then collaboratively, ensuring an exhaustive list of ideas (Figure 11). The (4) opportunity selection (Appendix 1.3.4) task was refined using a two-criteria power-dotting exercise (Desirability and Feasibility), ensuring equal influence for all participants regardless of their position or preference for extraversion. Finally, participants were tasked to form an alliance to perform the (5) opportunity framing task (Appendix 1.3.5). For this last task, an element of the new OM-session method was integrated, replacing a comparable task from the original design. This was changed to support the transfer of information from the Opportunity Identification (OI) session to the Opportunity Modeling (OM) session.

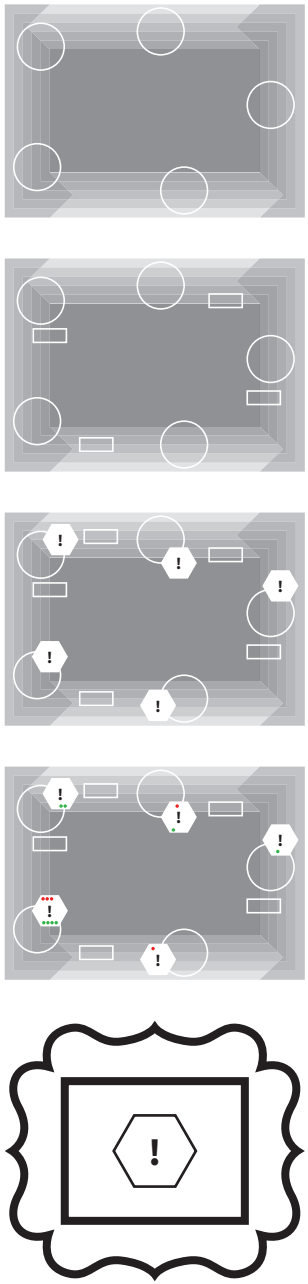


Figure 9 - OI session and task design

OPPORTUNITY SELECTION

Selecting the opportunities in the original design was a very minimal procedure, providing little time to properly evaluate the identified opportunities. Splitting the workshop offered the opportunity to perform a separate and expanded selection exercise, made it easier to involve key decision-makers, which was important since "effective learning from models occurs best, and perhaps only, when the decision-makers participate actively in the development of the model" (Sterman, 2000).

Additionally, Bingham (2014) stresses that a focused and disciplined selection process is instrumental in preventing cognitive dissonance and suggests the creation of an opportunity sequence or priority list. To "bring the present and future together in a way that facilitates team alignment and channels the energy and attention of geographically dispersed employees and managers". It allows individuals to take time to consider how opportunities can and should be ordered, helping them resist responding to any first potential, and allows more divergent thinking by considering the potential linkages between other trends and opportunities (Davis & Bingham, 2012).

Thus, for the design of the opportunity selection, a four-step process was designed (Figure 10) to be performed with a select group of decision makers who represent each synergy perspective (Appendix 1.4). After introducing the session (Appendix 1.4.1), (1) participants are tasked to group the identified opportunities and identify any sequences (Appendix 1.4.2). After this they are (2) tasked to map the opportunities on a two-by-two matrix, with a short/long term axis and a high/low priority axis (Appendix 1.4.3). Next, (3) the team is asked to select the most interesting or important opportunities by creating a priority list from bottom to top (Appendix 1.4.4). Lastly, (4) the team is asked to select an opportunity that they believe should be the focus of the participatory modeling workshop, then discuss which perspectives should be included, and then form a short list with potential representatives (Appendix 1.4.5).

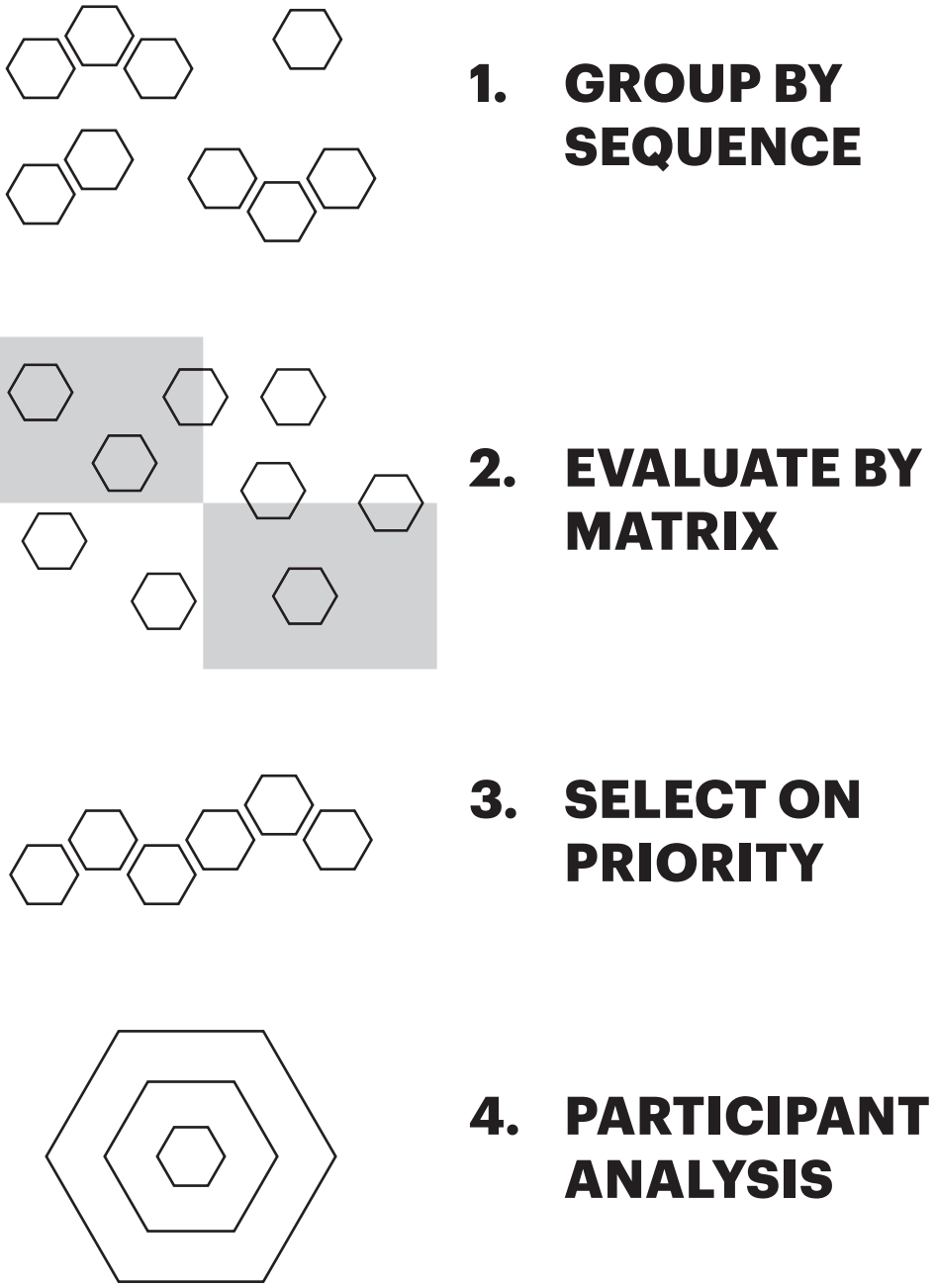


Figure 10 - OS session and task design

OPPORTUNITY MODELLING

Participants of the original group modeling process in the workshop were not able to successfully perform this task. The attempt of the Raaphorst (2023) study to incorporate the five BDM steps likely failed because it misinterpreted the use-case of the five-step BDM, using it in a qualitative setting while intended for a quantitative setting (Coyle, 2009; Kunc, 2017; Sterman, 2000).

A qualitative approach analyses the systems behavior with the help of a conceptual mental model (Coyle, 2009; Wolstenholme, 1999), resulting in a model that embodies "a set of assumptions describing a problematic situation" (Lane, 2000, p. 241).

Complexity & Quantitative vs Qualitative modelling

This study considers Organizational Synergy to be complex since it houses not only the system-induced complexity (Byrne & Callaghan, 2014; Jackson, 2019) of combining various dynamic organizational structures in which people and groups are viewed as simple components and agents. But it also embodies 'general complexity' (Morin, 2007); as explained by Byrne and Callaghan (2014) "individuals are themselves complex systems, and certainly more complex in every way than the agents in agent-based simulations" (Byrne & Callaghan, 2014, p. 41). It is about "nested and interpenetrating complex social systems beyond individuals, although of course with individuals as elements in those systems" (p. 41), "with causal powers running in all directions" (p. 45). Therefore, not all elements are quantifiable, making it difficult to quantify cause and effect links by equation. Instead, the designed synergy creation process must make do with soft, hard-to-measure elements from the organizational systems and socially influenced data obtained from interviews (Brychkov, Domegan, & McHugh, 2022).

Wolstenholme (1999) that the need for quantification in system modeling depends on objectives, methodology, and audience. This study assumes that, because its objective is complex, its methodology is qualitative, and it concerns a select group of individuals, it is justified to use a qualitative system modeling approach.

Figure 11 - Impression Use-Case OI-Session

OPPORTUNITY MODELLING

Group Model Building

To support a qualitative system modeling approach, a new method was selected, namely Group Model Building (GBM). The essence of GMB is that stakeholders are empowered to insert their expertise and modify a tangible representation of the dependencies across system dimensions (Black & Andersen, 2012) to exchange mental models, communicate key elements and causal links to elicit problem-affecting feedback loops (Sterman, 2000; Rees, 2017; Hovmand P. S., 2014).

GMB methods have been under development for several decades. Roberts (1977) had already addressed the importance of groups in the modeling practice, before a formal GMB definition was presented by Vennix (1996)- (1997). After which various authors (Richardson G. P., 1995; Richmond, 1997; Eden & Ackermann, 1998; Meadows, 1999; Ford, 1999; Cavana & Maani, 2000; Sterman, 2000; Howick, Ackermann, & Andersen, 2006; Andersen, et al., 2006; Wright & Meadows, 2008) (Richardson G. P., 2006; Hovmand P. S., 2014; Morecroft, 2015) contributed to the current state-of-the-art, which is captured in an up-to-date web-based library info, materials and validated methods and approaches, named Scriptapedia (Hovmand, et al., 2011).

From these materials, a toolkit was adopted from the System Mapping Academy to replace the original modeling process. It has the same objective as the Raaphorst (2023) study and is, in many ways, comparable to the core structure of the original Raaphorst (2023) design. It is, however, more refined, provides clear examples, and is more attuned to deal with a qualitative setting. The toolkit provides an openly accessible Miro template (Wick & Gampp, 2019) that covers a five-step process (Figure 12); (1) Framing the system, (2) Explore the system, (3) Map the system, (4) Reflect on the map, (5) Find leverage (Appendix 1.5).

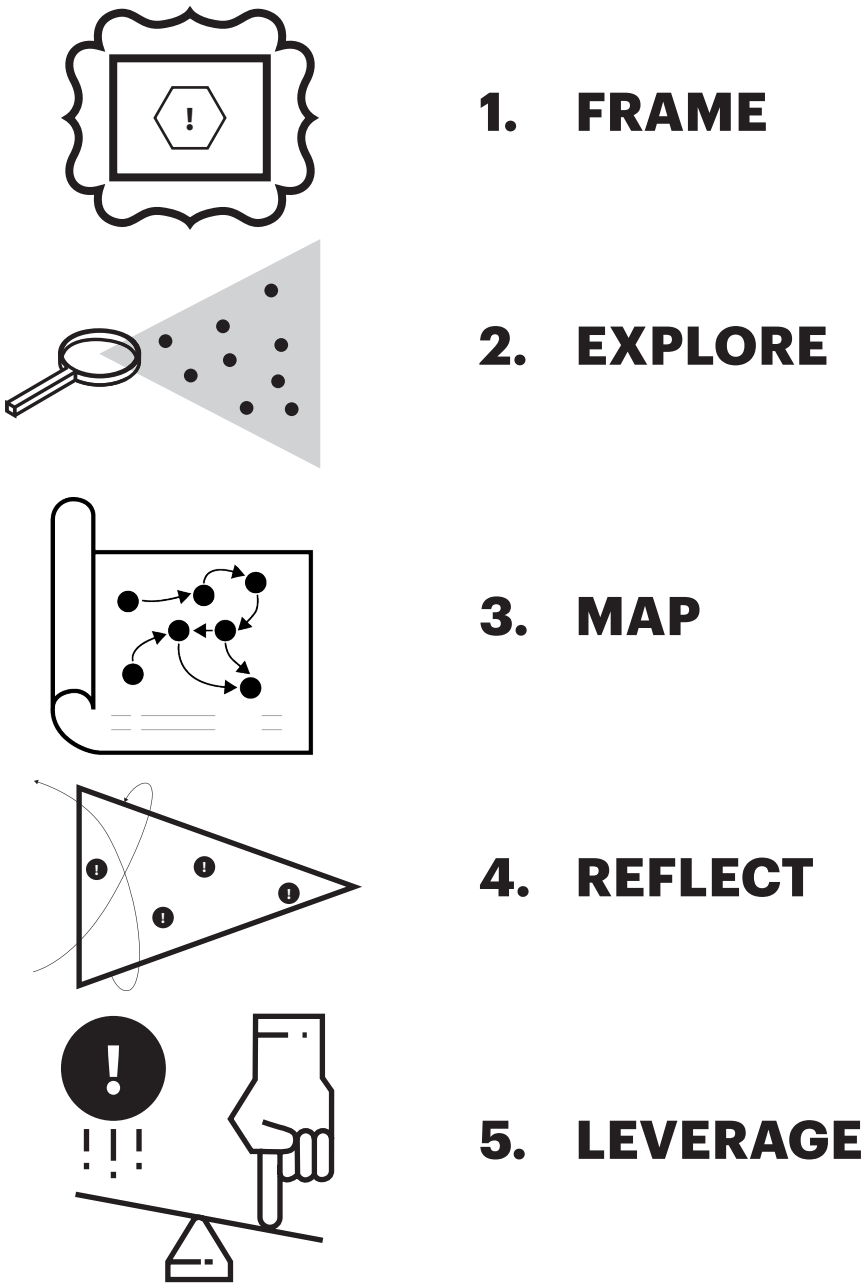
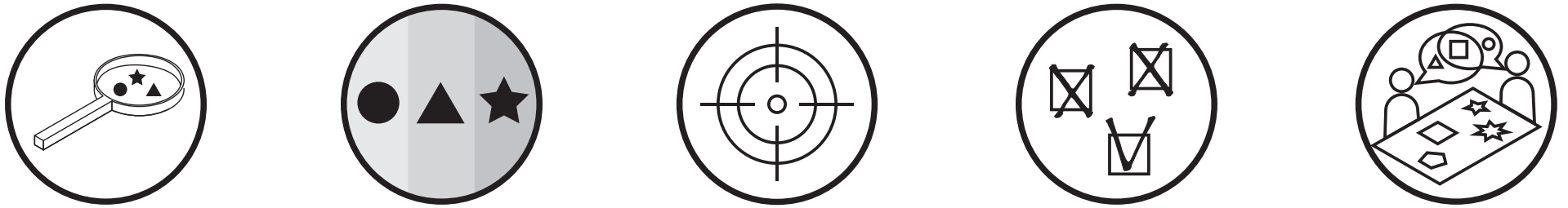


Figure 12 - OM session and task design

RE-DESIGNED SCP OVERVIEW

The three-phase process that involved 16 tasks was redesigned into a five-phase process that embodies 22 tasks (Figure 13).



CONTEXT CAPTURING	BLUEPRINT CREATION	OPPORTUNITY IDENTIFICATION	OPPORTUNITY SELECTION	OPPORTUNITY MODELLING
Business model Canvas	Word Count	Areas of interest	Group	Frame Opportunity
Customer Journey	Themes	Position Existing Synergies	Evaluate	Explore
Existing / Emerging Synergies	Chronological Project Order	Identify new potential synergies	Selection	Map
	Split in Groups	Pre-select by Power Dotting	Recruitment	Reflect
	Differences & Similarities	Form Alliance & Frame Opportunity		Leverage

Figure 13 - Overview of re-designed Synergy Creation Process (Sessions and Tasks)

RESULTS

Detailed description of study population and the resulting raw data

PILOT RESULTS

Firstly, the redesign of the SCP process was tested in a controlled environment (among interns) before deploying it in a real business case.

Conducting a pilot study allows for identifying potential issues and areas for improvement in a low-risk setting. It helps refine the process, mitigating flaws and improving efficiency.

For which the objective was formulated as 'exploring synergy between interns'. It was selected based on convenience, accessibility, and its potential to generate relevant output for the participants. This also allowed for evaluating the data collection method to prevent survey fatigue.

Full census was not achieved as not all approached participants were able to participate. (Table 4).

Session	Interns	Supervisors	Total
(CC) Context Capturing	6	4	10
(OI) Opportunity Identification	3	5	8
(OS) Opportunity Selection	6	3	9
(OM) Opportunity Modelling	2	2	4

Table 4 - Pilot sample size and group distribution per session

Figure 14 - Impression Pilot Study

DESIGN REFINEMENTS

Various areas of improvement were identified based on participant comments (Appendix 4.x.1), the TLX-comparison (Table 5) between the pilot and the benchmark (Appendix 4.x.2), the internal TS-comparison with corresponding participant feedback (Appendix 4.x.3), and the UEQ-comparison (Figure 15) to its native benchmark (Appendix 4.x.4). Various small design refinements were identified. The analysis and its resulting design considerations can be found in Appendix 4.6.

PILOT TASK LOAD	Pilot			Bench			P- Value	Effect Size
	Mean	SD	N	Mean	SD	N		
CC total	4,6	1,4	10	6,0	0,6	4	0,055	0,700
CC Starting	5,0	1,5	6	6,0	0,6	4	0,238	0,500
CC Follow-Up	3,9	1,1	4	6,0	0,6	4	0,029	1,000
OI	4,2	1,2	8	5,7	0,6	4	0,050	0,750
OI	5,1	0,9	9	5,7	0,6	4	0,435	0,306
OM	4,5	0,7	4	5,7	0,6	4	0,057	0,875

Table 5 - Combined Task Load statistics (Pilot vs Benchmark)

In summary, the most prominent design changes were:
1. Improving the overall process explanation to improve expectation management and session focus.

2. Refining the session briefs, template prompts and template structures, making them more explicit, by providing examples and clearer definitions of expected outcomes.

3. Switching the first and second tasks in the selection session to improve its flow.
4. Adding a sub-task to the modelling process to support variable prioritization, to make the start of the modelling task in the OM-session easier.

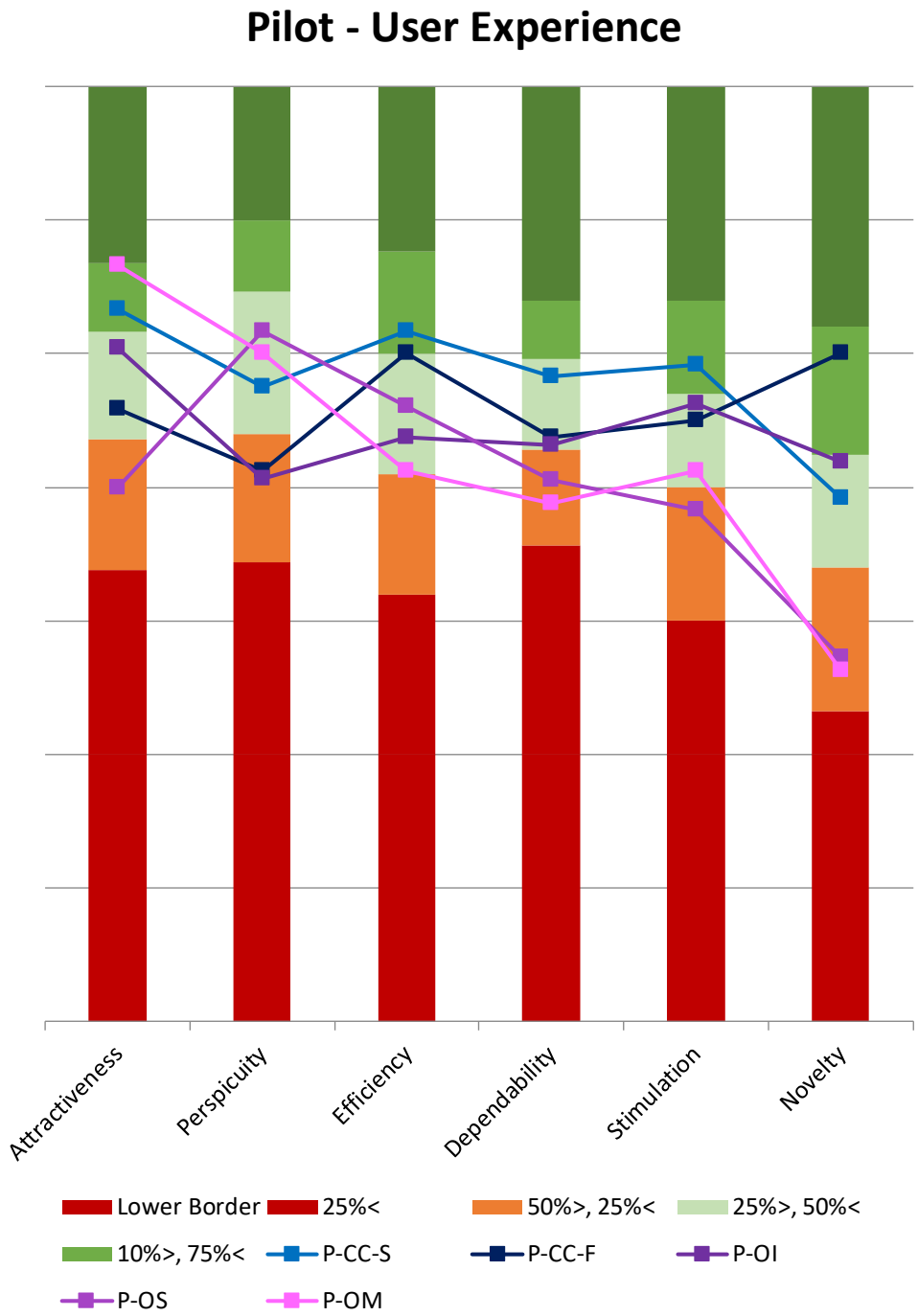


Figure 15 - User Experience all pilot sessions plotted on the Benchmark

USE-CASE RESULTS

Actual Sample Size

The invited and participating samples strongly differed in size, as the availability or willingness to participate was lower than expected (Overview in Table 6). With 15 participants, the CC session was the only one that exceeded the minimum sample size (Table 7). For the OI Session, due to unexpected circumstances (rescheduling of the originally planned date), only 10 participants were able to fully participate and complete the survey, resulting in a margin of error of 6,46% instead of the intended 5% (Table 8). Similarly, the required sample size was not met for the OS and the OM session. Resulting both in a sample size of four participants leading to a respective margin of error of 9,75% for OS (Table 9) and 10,22% for the OM session. (Table 10). Not meeting the required sample sizes for the OI, OS, and OM sessions means higher margins of error. Therefore, the results should be treated with caution and are seen in this study as indicative, not conclusive.

Session	H&PS	IX	ID	ESP	S&C	Total
(CC) Context Capturing	2	4	2	4	3	15
(OI) Opportunity Identification	1	2	2	2	3	10
(OS) Opportunity Selection	0	0 (1)	2	1	1	4
(OM) Opportunity Modelling	1	1	1	1	0 (1)	4

Table 6 - Use-case sample size and group distribution per session

POPULATION CC-SESSION LEVEL DISTRIBUTION						
LEVEL	H&PS	IX	ID	ESP	S&C	Total
11	0	0	0	1	1	2
10	0	0	0	1	0	1
9	3	2	0	1	0	6
8	0	0	0	1	0	1
7	1	1	2	2	1	7
6	1	1	1	1	1	5
5	0	1	0	0	0	1
4	0	0	0	0	0	0

Table 7 - Seniority Level Distribution CC-Session

POPULATION OI-SESSION LEVEL DISTRIBUTION						
LEVEL	H&PS	IX	ID	ESP	S&C	Total
11	1	0	0	0	2	3
10	0	0	0	0	0	0
9	0	1	0	0	1	2
8	0	0	0	0	0	0
7	0	0	2	1	0	3
6	0	1	0	1	0	2
5	0	0	0	0	0	0
4	0	0	0	0	0	0

Table 8 - Seniority Level Distribution OI-Session

POPULATION OS-SESSION LEVEL DISTRIBUTION						
LEVEL	H&PS	IX	ID	ESP	S&C	Total
11	0	0	0	0	0	0
10	0	0	0	0	0	0
9	0	0	0	0	0	0
8	0	0	0	0	0	0
7	0	0	2	0	0	2
6	0	0	0	1	1	2
5	0	0	0	0	0	0
4	0	0	0	0	0	0

Table 9- Seniority Level Distribution OS-Session

POPULATION OM-SESSION LEVEL DISTRIBUTION						
LEVEL	H&PS	IX	ID	ESP	S&C	Total
11	0	0	0	0	0	0
10	0	0	0	0	0	0
9	1	1	0	0	0	2
8	0	0	0	0	0	0
7	0	0	1	1	0	2
6	0	0	0	0	0	0
5	0	0	0	0	0	0
4	0	0	0	0	0	0

Table 10 - Seniority Level Distribution OM-Session

FINDINGS

Interpretation of the data per session

CROSS SESSION FINDINGS

To provide a clear analysis, each session is covered separately. For each session, the two research questions (RQ) are addressed. Starting with (TLX-RQ-1) 'How does the redesign of the SCP impact the task load of each individual session compared to the initial workshop design?' Followed by (UEQ-RQ-2) 'What is the user experience of participants per session in comparison to the standardized UEQ benchmark?'

For an overview of the results, please consult Appendix 5; the results and analysis are addressed in detail in Appendix 5.1 – 5.5. Before addressing the individual sessions, it is important to highlight three cross-session findings before analysis.

1. TLX - Reduced Task Load & Perceived Performance Decline
A notable observation from each session was the simultaneous reduction in perceived performance and TL indicators. This does not necessarily indicate a decline in actual performance, as it might also indicate that participants' perception of their performance is closely linked to the experienced TL. When TL is high, participants tend to perceive their performance as better, possibly because they feel more engaged or challenged.

2. TLX - Low Frustration Baseline
Another significant finding is the higher frustration levels compared to the benchmark. However, as explained in the Analysis Approach, the benchmark itself is set at a very low level (1 on a 10-point scale), which could make it a misleading comparison. The inherent low benchmark provides an extremely low baseline that is difficult to compete with, suggesting that its P-values and Effect sizes might not be as concerning as they initially appear.

3. TS - Task Specific Task Load Balance
During the pilot phase, task-specific metrics indicated imbalances in difficulty, temporal, and mental demand. After refining the sessions, no significant differences with substantial effect sizes were observed when comparing tasks internally. This finding was true for the CC (Appendix 5.1.3), OI (Appendix 5.2.3), OS (Appendix 5.3.3), and OM-session (Appendix 5.4.3). Based on this finding, this study, therefore concludes that the tasks were equally difficult, mentally, and temporally demanding, indicating a proper balance between tasks in every session.

CC-FINDINGS

The Starting Participants (Figure 16) with an empty template reported lower TL and better overall UX compared to those who reviewed their peers' templates (Figure 17). This effect is most strongly visible in the decrease in perspicuity and efficiency. This implies that integrating the Modified Delphi method had a slightly negative effect on TL and UX, possibly due to the increased amount of information to process. This was also reflected by written participant feedback (Appendix 5.1.1).

"It could help to hide the content and show it when you progress during the meeting, its overwhelming if you see it all."

Nonetheless, some participants appreciated building upon others' work, indicating the potential benefits of this method on the quality of the session output.

"Good to show input from others, can be built upon."

CC-Conclusion

TLX-RQ-1: The redesign of the SCP significantly reduced the TL compared to the benchmark. Small design changes aimed at improving communication likely contributed to this positive impact. However, the integration of the Modified Delphi method increased the TL slightly due to the need for participants to process more information. Despite this, the TL remained lower compared to the initial workshop design.

UEQ-RQ-2: The UX of the CC session compared to the general benchmark was generally positive, scoring high on attractive-ness and perspicuity, especially among participants who worked with empty templates. However, those who reviewed their peers' work reported lower scores on perspicuity and efficiency, likely due to the additional information they needed to process. Overall, the CC session was well-received, scoring above average on all UEQ metrics except for novelty.

*Circle Diagrams: The benchmark means were used as zero-line to visualize how the scores performed against the benchmark average. Scores on the right side of the line scored higher, and scores on the left were lower than the benchmark average.

CC-PERFORMANCE VS BENCHMARK (FIRST)

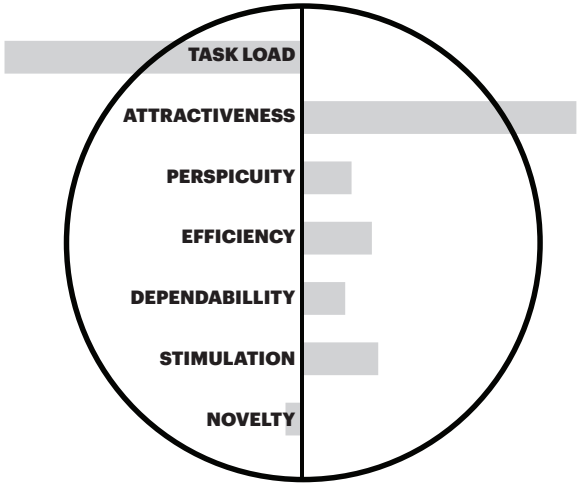


Figure 16 - Context Capturing interviews (First Participants)
[Compared to benchmark average (*p. 19)]

CC-PERFORMANCE VS BENCHMARK (FOLLOW-UP)

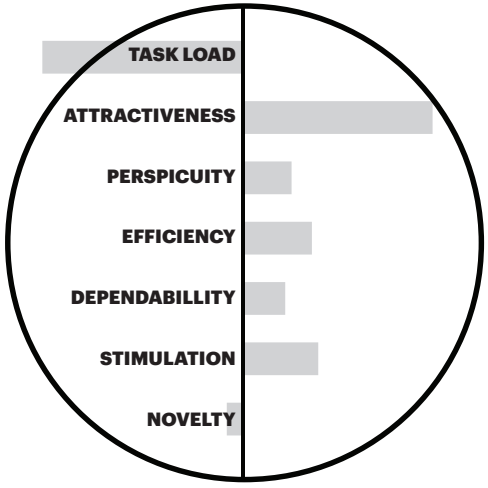


Figure 17 - Context Capturing interviews (Follow-up Participants)
[Compared to benchmark average (*p. 19)]

OI-FINDINGS

The OI session (Figure 18) showed a significant reduction in TL but retained a relatively higher perceived performance, which could indicate that the session is properly balanced while challenging. It however did also show a strong increase in Frustration, which might be attributed to the readability and information volume of the blueprint or because the session was a bit crowded, making it sometimes difficult to reach the blueprint (Appendix 5.2.1).

"Maybe less people or bigger roadmap"

Additionally, the session scored poorly on the Perspicuity UEQ benchmark, with participants finding it confusing and complicated. This was likely also due to the service blueprint being overwhelming. Next to this, the low Efficiency score suggests that the session's pace could be increased a bit. This is however contradicting with Perspicuity as participants require time to process information.

Also, Dependability scored below average, primarily on predictability. This might, however, be inherent to the session's exploratory goals and might therefore not be fully representative. It is suspected that the session's nature contributed to low predictability, a factor that may, in this context, not be problematic.

OI-Conclusion

TLX-RQ-1: The redesign significantly reduced the TL for the OI session. The new structure provided a balanced challenge, allowing participants to achieve high perceived performance with relatively low effort. This indicates the redesign's success in distributing the workload more effectively. Addressing accessibility and information volume challenges could further improve TL.

UEQ-RQ-2: The UX for the OI session scored both higher and lower than the benchmark. While the session scored well on attractiveness and stimulation, it performed poorly on Perspicuity and Efficiency. Some participants found the session confusing and complicated, likely due to the high volume of information on the service blueprint. Dependability scores were also low, primarily due to the session's inherent unpredictability. Despite these challenges, the session was perceived as valuable and interesting.

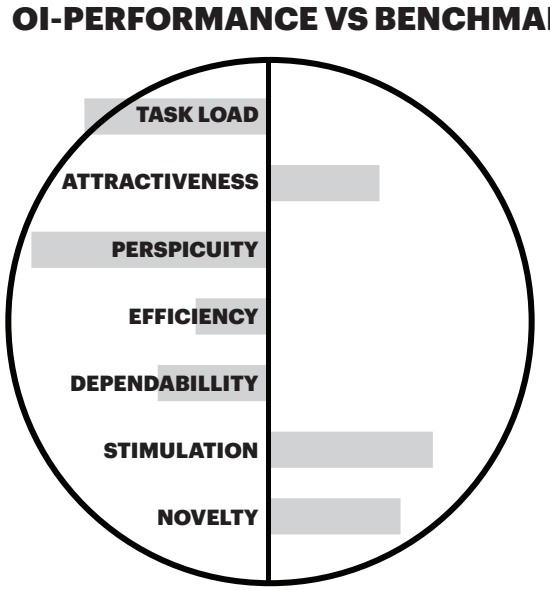


Figure 18 - Opportunity Identification session (Follow-up Participants) [Compared to benchmark average (*p. 19)]

OS-FINDINGS

Despite attempts to improve the OS-session, the revised design performed worse on all metrics compared to the pilot (Figure 19). Notably, perspicuity and efficiency saw severe drops and frustration was high, possibly due to higher expectations and more critical assessments from the senior participants involved in the use-case. For the OS-session four pain points were identified based on written participant comments (Appendix 5.3.1).

1. Opportunity Meaning Interpretation: Firstly, the interpretation of the typed-out opportunities was challenging, as they only provided a brief description that was detached from its context. A participant framed it as, *"it's just titles - making assumptions on what is meant, discussion time limited"*.

2. Axis Definition Confusion: Secondly, participants struggled to interpret the provided selection criteria (shot/long term & high/low priority) and felt the need to consider multiple definitions.
"During the session colleagues needed a definition of the axes. Most are used to setting the priorities and timeline based on other dimensions, such as value/effort. Urgent / important..."

3. Internal vs. External Focus: Thirdly, participants felt the need to differentiate between the internal and external focus of the opportunities, requiring much time and causing confusion.
"We had 'opportunities' that were related to internal, it was somewhat confusing on what was asked for."

4. Repetitive Actions: Lastly, the sequencing, matrix mapping and then prioritizing was considered as an inefficient flow, it "f*elt a bit double*".

The experienced unclarities and flow inefficiency led to a negative impact on the "*time to clarify the definitions of opportunities and for a proper comparison*", which was both observed by the facilitator as well as expressed by the participants.

OS-Conclusion

TLX-RQ-1: The TL for the OS-session showed a notable reduction except for an increase in frustration. Due to the sample size only, mental demand achieved significance. The reduction on Temporal Demand was quite a bit lower, this is likely due to various clarity and efficiency issues that might have impacted the sense of available time.

UEQ-RQ-2: The UX for the OS session was the lowest among all sessions. Several issues arose from the redesign. Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty all scored far below average. Participants found the session confusing, inefficient, and repetitive. The cause for this was the loss of context in the transfer of information from the OI-session and unclarity about session objectives and selection criteria. Thus, substantial revision is required.

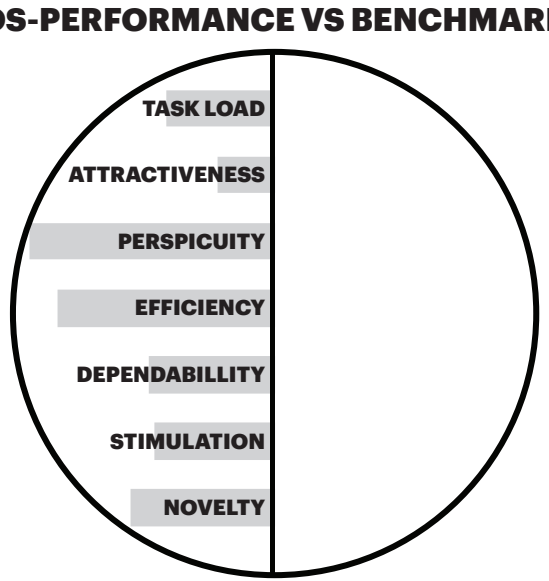


Figure 19 - Opportunity Selection session (Follow-up Participants)
[Compared to benchmark average (*p. 19)]

OM-FINDINGS

Despite the need for halving the session duration, which likely impacted mental and temporal demand, the Group Modeling method showed a TL reduction in effort and perceived performance (Figure 20). Additionally, the session scored highly on stimulation and novelty, indicating its value and interest to participants. Some participants, however, found the modeling method overcomplicated and indicated that not all topics might be suitable for this method. For simpler topics, the method might seem excessive, while for complex ones, more time and workforce are needed to produce reliable results.

The main identified area of improvement was the implementation of the session. This was mostly based on the low Perspicuity, Efficiency, and Dependability score, which suggests that improving pre-session communication and more extensively preparing participants could enhance these UX metrics. Improving these metrics could potentially also further reduce the mental and temporal demand of the session. This improvement was also indicated via participant comments (Appendix 5.4.1):

"Perhaps a more detailed introduction of the process and exercise at hand, could help with a faster thought process?!"

When considering the Perspicuity, Efficiency and Dependability of all other sessions, a trend emerges that suggests that the other sessions would also benefit from improving pre-session communications (Figure 21).

OM-Conclusion

TLX-RQ-1: The OM-session showed a reduction in TL, particularly in effort and perceived performance. However, high mental and temporal demands remained due to an unexpected, shortened session duration. The TL improvements were less pronounced compared to other sessions, likely due to the complexity and inherent difficulty of the session's objective.

UEQ-RQ-2: UX for the OM session was mixed. While it scored high on stimulation and novelty, it performed poorly on perspicuity, dependability, and efficiency. Some participants found the session confusing and slow. Additionally, a need was found for better pre-session communication and preparation. Despite these challenges, the session's innovative and engaging nature was appreciated.

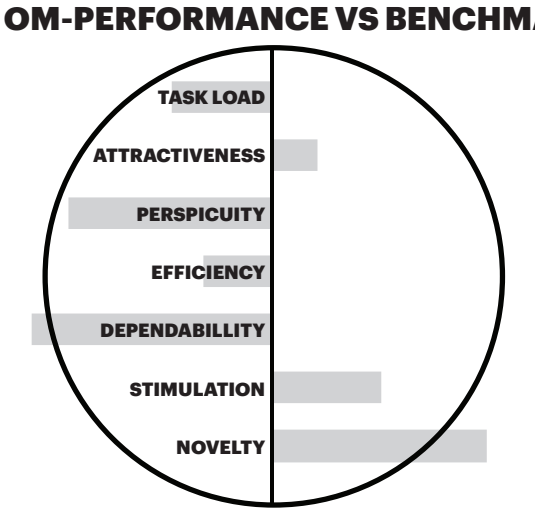


Figure 20 - Opportunity Modelling session (Follow-up Participants) [Compared to benchmark average (*p. 19)]

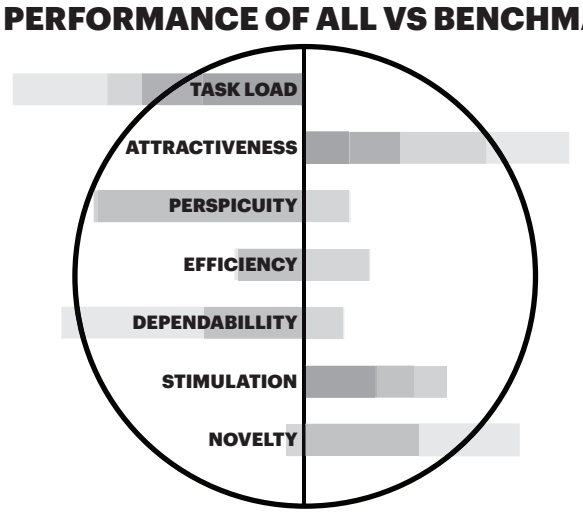


Figure 21 - All sessions combined (Follow-up Participants) [Compared to benchmark average (*p. 19)]

SCP - EFFECTIVENESS

No direct evaluation of the SCP's effectiveness was conducted. However, some items of the UEQ questionnaire could provide some insight into this process's perceived value.

The CC and OM session, for instance, scored higher than 90%, and the OI session above 75% of the stimulation benchmark for the **inferior/valuable** item. The CC and OI sessions both scored above 90% and OM above 75% of the stimulation benchmark on the **not interesting/interesting** item (Appendix 5.5).

The scores of these items do not directly indicate effectiveness but do suggest that these phases were perceived as valuable and interesting.

CONCLUSION

The answer to the research questions per session

CONCLUSION

So, did the redesign reduce the task load? Yes, the redesign of the SCP successfully reduced the TL across all sessions compared to the initial workshop design. However, each session faced specific challenges. The CC session saw a slight increase in TL due to the Modified Delphi method, while the OS session struggled with inefficiencies in information transfer and task clarity. The OM session, despite a reduction in effort, continued to experience high mental and temporal demands.

And how well did the UX of the session perform against the UEQ benchmark? Well, the UX varied significantly across sessions. The CC session had a high UX, especially for participants using empty templates. While engaging, the OI session was rated poorly in terms of clarity and efficiency. The OS session performed the worst, with low scores on all UX metrics. The OM session had mixed results, scoring high on stimulation and novelty but low on clarity and efficiency.

In conclusion, the SCP seems to provide us with a new, fairly viable (based on TL), and relatively user-friendly approach to explore organizational synergy in a valuable, interesting, and constructive way. And however the design requires further development (mostly on perspicuity, efficiency, and dependability), it yields promise for supporting complex merger and acquisition processes.

DISCUSSION

Since various design elements have been implemented, and because the SCP was tested in a slightly different context while incorporating more perspectives, a conservative approach to the interpretation is appropriate.

We can segment the design changes into four parts: (1) the Modified Delphi method, (2) the Workshop Split, (3) the Expanded Selection activities, and (4) the Group Modelling method. These changes, except for the split, are isolated and can therefore be considered separately. The workshop split, however, resulted in the three OI, OS, and OM-sessions, thus impacting all three of these sessions.

Additionally, the pilot test also resulted in various minor design changes (like balancing the TL per session, clarifying explanations, and improving the session flow) across all the four interactive sessions. It needs to be considered that these changes have impacted the results for all sessions.

When we consider each individual design change we find a different effect on the TL and UX. The **Modified Delphi Method** did not prove fruitful and appeared to have a slight negative effect on both TL and UX. Some participants however expressed appreciation for the approach, indicating the need for further investigation to assess its impact on the quality of the output and whether this outweighs the minor negative effects on TL and UX.

Splitting the workshop significantly reduced the TL for all three segments compared to the benchmark workshop. However, it also equally decreased participants' perceived performance. While the design change achieved its intended objective, it needs to be considered if a proper balance between TL and perceived performance is necessary for an effective session and how this can be achieved.

During the **Expanded Selection** session, various challenges emerged. Due to a loss of context when transferring information between sessions, participants had to make assumptions about the meaning of the developed opportunities. Additionally, it performed poorly on UX due to unclear objectives, too broad selection criteria, and an inefficient flow of tasks. These challenges likely also limited the reduction of temporal demand that was induced by the split.

The **Group Modeling** session was cut short due to unexpected circumstances, likely contributing to higher temporal demand. Although the session showed significant effort reduction, it still recorded high mental demand. Additionally, it scored low on Efficiency, Dependability, and Perspicuity and surprisingly high on Stimulation and Nov-

elty. Therefore, whether Group Modeling is a good method for supporting organizational synergy remains inconclusive. Implementing the method differently by improving pre-session communications, providing more perspicuity, and dedicating more time might mitigate challenges in Efficiency and Dependability.

Lastly, when comparing the results of the Pilot and the use-case it stands-out that the Pilot performed quite a bit better than the use-case. This might be due to two reasons: (1) the intern projects were inherently less complex by nature, and (2) more senior employees are more critical and have a higher standard. Since the benchmark was measured in a similar context to the use case, a comparison to the Pilot is considered unreliable for drawing conclusions about the effect of the redesign on the TL and UX.

In summary, the workshop split and the minor overall improvements based on prior recommendations and the pilot findings, seemed to have enhanced the overall TL. The Modified Delphi and Expanded Selection activities showed no such improvements and require further development and research. The Group Modeling method showed promise, but its implementation could be improved and should be tested again, as this test was compromised by a shortened session duration.

LIMITATIONS

Initially, it was the ambition to perform multiple cycles that would allow for a stronger comparison and further design refinements. A major limitation of this research was, however, the availability of participants, causing severe time constraints and impacting intended sample sizes, resulting in larger margins of error (5-10%). Therefore, only the pilot and one evaluation cycle could be completed. For this reason, this study had to rely on the TLX and UEQ benchmark comparison and indicative output data.

The selected standardized questionnaires, SEQ, TLX, and the UEQ, are widely used methods. However, they rely on subjective self-reporting and are, therefore, not objective measures. They capture a snapshot in time and may not account for changes in perceived workload or user experience over time. In addition, the translation and interpretation of the survey items could have been interpreted differently due to the multi-cultural, multilingual population. In part, this was attempted to be mitigated by providing both validated Dutch and English translations in the survey.

In addition, the UEQ is designed for broad application across different products and services. This generality means it may not capture the nuances or provide a comparable benchmark relevant to the specific environment of the Synergy Creation Process (SCP) in a corporate setting.

The TLX lacks a validated benchmark. The study, therefore, relied on comparing the to the antecedent study, which are not validated and lack a substantive sample size. causing the data to have more extreme outliers (e.g a one out of ten ‘Frustration’ mean score)

It also must be noted that the Post-Task and TLX survey itself might have impacted the cognitive load experienced by the participants. The expertise level, overall mood, and fatigue of the participants outside of the workshop might also have influenced the perceived workload.

Lastly, since the Usability sub-scale results were deemed unreliable. This research only evaluates the design changes on how nice it is to perform the process as a participant. While valuable insights for improvement, it does not indicate how useful, effective or qualitative the output is.

FUTURE WORKS

Developing a process that can be deployed to drive organizational synergy could provide a powerful tool for initiating innovations and improving businesses. The process as designed for this thesis has been improved on various points; however, it requires further development and research, as for many parts of the process, challenges and pain points were identified.

Most of all, it is recommended that research be continued on the effect of the generated output on the synergy efforts of the organizations. Various considerations have been discussed in this paper that could influence the quality of the output, such as the potential positive effect of the Modified Delphi method on the data collection.

Next to this the TL of the facilitator needs further improvement. The creation of the service blueprint was placed out of scope for this study. It however required considerable time to create, making the deployment of this process highly time-consuming for the facilitator.

Further, this study identified various pain points across the process that would benefit from a design revision. Primarily, the OS-Session has shown much room for improvement, but the Perspicuity, Efficiency, and Dependability of all three workshop sessions can be improved. A recommendation for improving these sub-scales is the creation of a communication protocol, the data indicating a lack of understanding of the session objectives. Clearer and more intensive briefings and pre-session communications are likely to improve all these three metrics. Improving objective communication could also help in selecting a suitable topic for the Group Modelling workshop. Rendering the process more effective and targeted towards the needs of the organization.

After improving the process, to the point that both UX and quality of output have been optimized, it would be highly interesting to perform behavior change research to understand the effect of this process on the short and ideally also long-term behavior of participants and the organizations.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my university mentor, Lu Yuan, for her outstanding support during this study. Her guidance was spot on, providing me with metaphorical fishing lessons instead of offering a fish. Secondly, I would like to thank my Accenture supervisors, Laetitia van Wijnen and Rens van Mens, who opened their network to me, provided me with ample space to grow and helped me develop my project management skills. I also want to express a special thanks to Rodolfo Rangel, Niels van Maanen, Phillip Drift, Mark van der Laar, and Ellen Kageler for making time and sharing their thoughts. I also wish to express my gratitude to the VanBerlo team for opening their hearts to me once more, The Industry-X team for the warm welcome, and the ESP, S&P, and H&PS teams for participating enthusiastically during my journey.

AI-Tools

GPT-4 & GPT-4o were used during this study. It was used to reflect on the research protocol and process, identify literature, and support sentence phrasing and troubleshooting for coding during data processing. It has explicitly not been used for making design choices, data analysis, data interpretation and the generation of large text chunks. All AI output was reviewed and edited before integrating into this thesis.

REFERENCES

Accenture. (2024). Accenture reports Third-Quarter Fiscal 2024 results. Retrieved from https://newsroom.accenture.com/news/2024/accenture-reports-third-quarter-fiscal-2024-results

Accenture. (2024). Accenture, Netherlands, Let There Be Change. Retrieved from https://www.accenture.com/nl-en.

Accenture Newsroom. (2020, Februari 18). Accenture Acquires Dutch Product Design and Innovation Agency VanBerlo to Help Clients Build Smart Connected S. Retrieved from Accenture Newsroom: https://newsroom.accenture.com/news/accenture-acquires-dutch-product-design-and-innovation-

AgeLab, M. (2020). Three Models of Technology Adoption: A Literature Review in Brief . MIT AGELAB - CareHive.

Andersen, D. F., Bryson, J. M., Richardson, G. P., Ackermann, F., Eden, C., & Charles Finn, B. (2006). Integrating modes of systems thinking into strategic planning education and practice: the thinking persons' institute approach. Journal of Public Affairs Education, 265-293.

Argyris, C. (1985). Strategy, change and defensive routines. . Pitman Publishing.

Aronson, E. (1969). The Theory of Cognitive Dissonance: A Current Perspective. Advances in Experimental Social Psychology, vol. 4, ed. L. Berkowitz.

Atuahene-Gima, K. (2005). Resolving the capability–rigidity paradox in new product innovation. Journal of marketing, 69(4), 61-83.
Bauer, T. N., & Green, S. G. (1994). Effect of newcomer involvement in work-related activities: a longitudinal study of socialization. Journal of applied psychology, 79(2), 211.

Baumgartner, J., & Sonderegger, A. (2019). No need to read: Developing a pictorial single-item scale for measuring perceived usability. Sauer, J. . Int. J. Hum. Comput. Stud, 122, 78–89.
Beckhard, R., & Harris, R. T. (1987). Managing complex change. . Harvard Business Review, 65(3).

Bell, S., Kong, J., Clark, D., Carne, P., Skinner, S., Pillinger, S., . . . Brown, W. (2022). The National Aeronautics and Space Administration-task load index: NASA-TLX: evaluation of its use in surgery. ANZ J Surg, 92(11):3022-3028. doi:doi: 10.1111/ans.17830

Benner, M. J., & Tushman, M. L. (2003). Exploitation, exploration, and process management: The productivity dilemma revisited. Academy of management review, 28(2), 238-256.

Bingham, C. B., Furr, N. R., & Eisenhardt, K. M. (2014). The opportunity paradox. MIT Sloan Management Review, 56(1), 29.

Black, L. J., & Andersen, D. F. (2012). Using visual representations as boundary objects to resolve conflict in collaborative modelbuilding approaches. Systems Research and Behavioral Science, 29(2), 194-208.

Bos, C., & Lundberg, F. (2019). "A String of Pearls" - Exploring Organizational Motives and Challenges of Creative Agency Acquisitions by Leading Management Consulting Firms.

Boucher, W. I. (1980). The process of conglomerate mergers. Federal Trade Commission.

Brand, R., & Rocchi, S. (2011). Rethinking value in a changing landscape. A model for strategic reflection and business transformation. A philips design paper.

Braun, V., & Clarke, V. (2012). Thematic analysis. American Psychological Association.

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189.194 4-7.

Brown, T. (2008). Design thinking. Harvard business review, 86(6), 84.

Bruner II, G. C., & Kumar, A. (2005). Explaining consumer acceptance of handheld Internet devices. Journal of business research, 58(5), 553-558.

Brychkov, D., Domegan, C., & McHugh, P. (2022). Coming and going in loops: Participatory modelling of a system with all its complexity. Journal of Macromarketing, 42(1), 12-29.

Buchanan, R. (1992). Wicked Problems Thinking.

Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. Journal of Experimental Psychology: Applied, 13(4), 273.

Byrne, D., & Callaghan, G. (2014). Complexity theory and the social sciences: The state of the art.

Cavana, R., & Maani, K. (2000). A methodological framework for integrating systems thinking and system dynamics. In Proceedings of the 18th International Conference of the System Dynamics Society, pp. 6-10.

Chesbrough, H. W. (2003). Open innovation: The new imperative for creating and profiting from technology. Harvard Business Press.

Colligan, L., Potts, H. W., Finn, C. T., & Sinkin, R. A. (2015). Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. International journal of medical informatics, 84(7).

Coyle, G. (2009). Qualitative and quantitative modelling in system dynamics. System Dynamics, 2, 33.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. psychometrika, 16(3), 297-334.

Cross, N. (2011). Design Thinking: Understanding How Designers Think and Work. New York: Berg.

Cummings, T. G., & Worley, C. G. (2016). Organization development & change.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. Academy of management review, 32(2), 480-499.

Davis, J., & Bingham, C. (2012). Learning Sequences: Their Existence, Effect, and Evolution. Academy of Management Journal 55, no. 3, 611-641.

De Matteis, P. M. (2016). Why big consultancies buy agencies. Retrieved from https://www.campaignasia.com/article/why-big-consultancies-buy-design-agencies/407973

Díaz-Oreiro, I., López, G., Quesada, L., & Guerrero, L. A. (2019). Standardized questionnaires for user experience evaluation: A systematic literature review. MDPI, In Proceedings (Vol. 31, No. 1, p. 14).
Dorst, K. (2015). Frame innovation: Create new thinking by design. MIT press.

Dyer, W. G., & Dyer, J. H. (1994). Team building: current issues and new alternatives. http://ci.nii.ac.jp/ncid/BA26782833.

Eden, C., & Ackermann, F. (1998). Making Strategy: The Journey of Strategic Management. London: SAGE.

Fantaguzzi, I., & Handscomb, C. (2024, February). The importance of cultural integration in M&A: The path to success. Retrieved from mckinsey.com: https://www.mckinsey.com/industries/oil-and-gas/our-insights/the-importance-of-cultural-integration-in-m-and-a-the-path-to-success#/

Feagin, J. R., Orum, A. M., & Sjoberg, G. (2016). A case for the case study. UNC Press Books.

Fischer, T., Gebauer, H., Gregory, M., Ren, G., & Fleisch, E. (2010). Exploitation or exploration in service business development? Insights from a dynamic capabilities perspective. Journal of Service Management, 21(5), 591-624.

Ford, J. D. (1999). Organizational change as shifting conversations. Journal of Organizational Change Management, 12(6), 480-500.

Forster, Y., Hergeth, S., Naujoks, F., & Krems, J. (2018). How Usability Can Save the Day - Methodological Considerations for Making Automated Driving a Success Story. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, pp. 278–290.

Gianatasio, D. (2017). Global Consultancies Are Buying Up Agencies and Reshaping the Brand Marketing World. Retrieved April 17, 2019, from https://www.adweek.com/brand-marketing/globalconsultancies-are-buying-up-agencies-and-reshaping-the-brand-marketing-world/

Gupta, A., Smith, K., & Shalley, C. (2006). The interplay between exploration and exploitation. Acad.

Gustafson, D. H., Shukla, R. K., Delbecq, A., & Walster, G. W. (1973). A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups, and nominal groups. 9(2), 280-291: Organizational behavior and human performance.

Hannan, M., & Freeman, J. (1984). Structural inertia and organizational change. Sociological Review, 49, 149-164.

Hart, S. G. (1986). NASA task load index (TLX).

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. Sage CA: Los Angeles, CA: Sage publications. In Proceedings of the human factors and ergonomics society annual meeting, Vol. 50, No. 9, pp. 904-908.

Hassan, M. (2023). Purposive Sampling - Methods, Types and Examples. https://researchmethod.net/purposive-sampling/.

Hassan, Z. A., Schattner, P., & Mazza, D. (2006). Doing a pilot study: why is it essential? Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia, 1(2-3), 70.

Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13(4), 481-499.

He, Z. L., & Wong, P. K. (2004). Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis. Organization science, 15(4), 481-494.

Helfat, C., Finkelstein, S., Mitchell, W., Peteraf, M., Singh, H., Teece, D., & Winter, S. (2007). Dynamic Capabilities: Understanding Strategic Change in Organizations. Oxford: Blackwell Publishing.

Hovmand, P. S. (2014). Community Based System Dynamics. New York, NY: Springer Science, Business Media.

Hovmand, P., Rouwette, E. A., Andersen, D., Richardson, G., Calhoun, A., Rux, K., & Hower, T. (2011). Scriptapedia: a handbook of scripts for developing structured group model building sessions.

Howick, S., Ackermann, F., & Andersen, D. F. (2006). Linking event thinking with structural thinking: methods to improve client value in projects. System Dynamics Review, 22 (2):113-140.

Jackson, M. C. (2019). Critical systems thinking and the management of complexity. John Wiley & Sons.

Jansen, J. J., Van Den Bosch, F. A., & Volberda, H. W. (2006). Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. . Management science, 52(11), 1661-1674.

Jones, P. &. (2017). Rendering systems visible for design: Synthesis maps as constructivist design narratives. She Ji: The Journal of Design, Economics, and Innovation, 3(3), 229-248.

Kageler, E. (2024). Unlocking Potential: Business development for Accenture Life Sciences. Leiden Univeristy.
Kapoor, K. K. (2014). Conceptualising the role of innovation: Attributes for examining consumer adoption of mobile innovations. The Marketing Review, 14(4), 405-428.

Kautz, K. &. (2004). Understanding the implementation of software process improvement innovations in software organizations. Information Systems Journal, 14(1), 3-22.

Kelley, T. &. (2001). The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm. New York: Currency.

Klammer, J., & van den Anker, F. (2018). Design, User Experience, and Usability: Users, Contexts and Case Studies. Springer International Publishing: Berlin/Heidelberg, Germany, Volume 10920.

Krippendorff, K. (2006). Meaning, Participation, and Dialogue. Participatory Design Conference.

Kunc, M. (2017). System dynamics: A soft and hard approach to modelling. In 2017 Winter Simulation Conference (WSC), pp. 597-606. EEE.

allemand, C., & Koenig, V. (2017). How Could an Intranet Be Like a Friend to Me?: Why Standardized UX Scales Don'T Always Fit In. In Proceedings of the European Conference on Cognitive Ergonomics 2017 Umea, Sweden, pp. 9–16.

Lane, D. C. (2000). Should system dynamics be described as a 'hard'or 'deterministic'systems approach? Systems research and behavioral science: the official journal of the international federation for systems research, 17(1), 3-22 (p. 241).

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian.

Lavie, D., Stettner, U., & Tushman, M. L. (2010). Exploration and exploitation within and across organizations. Academy of Management annals, 4(1), 109-155.
Lawson, B. (2005). How Designers Think. New York: Architectural Press.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. . International Journal of Human-Computer Interaction, 7(1), 57-78.

Likert, R. (1932). A technique for the measurement of attitudes. Archives of psychology.

Linstone, H. A., & Turoff, M. (2002). Delphi bibliography. Journal Articles, 14(54), 68.

Lune, H. &. (2017). Qualitative research methods for the social sciences. Pearson.

Magistretti, S., Dell'Era, C., & Doppio, N. l. (2020). Design Sprint for SMEs: An Organizational Taxonomy based on Configuration Theory. Management Decision, 58 (9): 1803– 1817.

March, J. (1991). Exploration and exploitation in organizational learning. Organization Science 2:71–87.

Martin, R. L. (2009). The design of business: Why design thinking is the next competitive advantage. Harvard Business Press.

Martin, R. L. (2009). The design of business: Why design thinking is the next competitive advantage. Harvard Business Press.

Martin, R. L. (2009). The Design of Business: Why Design Thinking is the Next Competitive Advantage. Cambridge. MA: Harvard Business Press.

McKinsey & Company. (2010). A McKinsey perspective on creating transformational value from mergers. Retrieved from https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/a-mckinsey-perspective-on-creating-transformational-value-from-mergers

McKinsey&Company. (2018, februari 5). How the implementation of organizational change is evolving. Retrieved from mckinsey: https://www.mckinsey.com/capabilities/implementation/our-insights/how-the-implementation-of-organizational-change-is-evolving Meadows, D. (1999). Leverage points. Places to Intervene in a System, 19, 28.

Michael, D. N. (1997). Learning to Plan-and Planning to Learn. Miles River Press.

Miro. (n.d.). De visuele werkruimte voor innovatie. Retrieved from https://miro.com/. https://miro.com/nl/login/

Morecroft, J. D. (2015). Strategic modelling and business dynamics: A feedback systems approach. John Wiley & Sons.

Morin, E. (2007). Restricted Complexity, General Complexity. In Worldviews, science and us: Philosophy and complexity, pp. 5-29.

Nisbett, L. R. (1991). The Person and the Situation: Perspectives of Social Psychology. Philadelphia: Temple University Press.

Norman, G. R. (2010). Likert scales, levels of measurement and the “laws” of statistics. . Advances in Health Sciences Education, 15(5), 625–632. doi:https://doi.org/10.1007/s10459-010-9222-y

O’Reilly III, C. A. (2013). Organizational ambidexterity: Past, present, and future. Academy of management Perspectives, 27(4), 324-338.

Osterwalder, A. (2008). What is a business model? business-model-design.blogspot.com.

Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2015). Value proposition design: How to create products and services customers want. John Wiley & Sons.

Patrício, V., Costa, R. L., Pereira, L., & António, N. (2022). Dynamic capabilities and project management: a systematic literature review. International Journal of Business Innovation and Research, 29(4), 417-448.

Porac, J. a. (1990). Taxonomic mental models in competitor definition. Academy of Management Review, 15, 224-240.

Porac, J. F. (1990). Taxonomic mental models in competitor definition. Academy of management Review, 15(2), 224-240. .

Quinlan, R. (2018). Consultancies Are Increasingly Interested in Acquiring Marketing Agencies. Retrieved April 17, 2019, from https://www.consultancy.uk/news/18137/consultancies-are-increasinglyinterested-in-acquiring-marketing-agencies

Raaphorst, J. (2023). Project Harmony. M21_DIDM210.

Raisch, S., & Birkinshaw, J. (2008). Organizational ambidexterity: Antecedents, outcomes, and moderators. Journal of management, 34(3), 375-409.

Rees, D. J. (2017). Exploring Stability and Change in Transport Systems: Combining Delphi and System Dynamics Approaches. Transportation, 44 (4), 789-805.

Reichheld, F. (2011). The ultimate question 2.0 (revised and expanded edition): How net promoter companies thrive in a customer-driven world. Harvard Business Review Press.

Reus, M. d. (2023). Unlocking Synergy Potential - Designing a synergy creation approach for Accenture Industry X and VanBerlo. TU Delft.
Richardson, G. P. (1995). Teamwork in group model building. System Dynamics Review, 11 (2):113-137.

Richardson, G. P. (2006). Concept models. In Proceedings of the 24th International Conference of the System Dynamics Society.

Richmond, B. (1997). The strategic forum: aligning objective, strategy, and process. System Dynamics Review , 13 (2):131-148.

Rittel, H. W. (1973). Dilemmas in a General Theory of Planning. Policy Sciences, 4(2), 155-169. doi: https://doi.org/10.1007/BF01405730

Roberts, E. B. (1977). Strategies for effective implementation of complex corporate models.

Romme, A. G. (2004). Unanimity rule and organizational decision making: A simulation model. Organization Science, 15(6), 704-718.

Romme, A. G. (2010). Dynamic capabilities, deliberate learning and environmental dynamism: a simulation model. Industrial and Corporate Change, 19(4), 1271-1299.

Rubio, S. D. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile. Applied Psychology: An International Review, 53(1), 61-86. .
Rummler, G. A., & Brache, A. P. (2012). Improving performance: How to manage the white space on the organization chart. John Wiley & Sons.

Sample-Size-Calculator. (z.d.). Retrieved from https://www.calculator.net/sample-size-calculator.html?type=1&cl=95&ci=5&pp=1&ps=4000&x=Calculate

Sauro, J. &. (2009). Comparison of three one-question, post-task usability questionnaires. In Proceedings of the SIGCHI conference on human factors in computing systems , pp. 1599-1608.

Sauro, J., & Lewis, J. R. (2016). Quantifying the user experience: Practical statistics for user research. Morgan Kaufmann.

Schein, E. H. (1987). The clinical perspective in fieldwork. Sage Publications, Inc.

Schein, E. H. (1988). Organizational socialization and the profession of management. MIT Sloan management review, 30(1), 53.

Schön, D. &. (1996). Organizational learning II: Theory, method and practice. Addison Wesley, 305(2), 107-120.

Schon, D. A. (1983). The reflective practicioner: How professionals think in action . New York: Basic Books., p. 1983.

Schon, D. A. (1983). The Reflective Practitioner: How Professionals Think in Action. New York: Basic Books.

Schrepp, M. H. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108.

Schultz, E. J. (2019). Creeping on Agencies' Turf, Consultancies Are Shaking Up the Marketing Industry. Retrieved from adage: https://adage.com/article/news/consultancies-rising/308845/

Scott, R. J. (2016). Recent evidence on the effectiveness of group model building. European Journal of Operational Research, 249(3), 908-918.

Sekhar Chanda, S. R. (2018). The continuum conception of exploration and exploitation: an update to March's theory. M@ n@ gement, 21(3), 1032-1079.

Sevaldson, B. (2013). Systems Oriented Design: The emergence and development of a designerly approach to address complexity.

Sevaldson, B. (2018). Visualizing Complex Design: The Evolution of Gigamaps. Translational Systems Sciences, 243–269. doi:-doi:10.1007/978-4-431-55639-8_8

Simon, H. A. (1969). The Sciences of the Artificial. Cambridge: MIT Press.

Simsek, Z., Heavey, C., Veiga, J. F., & Souder, D. (2009). A typology for aligning organizational ambidexterity's conceptualizations, antecedents, and outcomes. Journal of management studies, 46(5), 864-894. .

Sköldberg-Johansson, U., Woodilla, J., & Çetinkaya, M. (2013). Design Thinking: Past, Present and Possible Futures. Creativity and Innovation Management, 22 (2): 121– 46.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of business research, 104, 333-339.
Stenberg, L. (1980). A modeling procedure for public policy. In Elements of the System Dynamics Method, edited by J. Randers. Cambridge, MA: MIT Press.

Sterman, J. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World. New York: McGraw Hill.
Stickdorn, M. H., M. E., L. A., & Schneider, J. (2018). This is service design doing. O'Reilly Media, Inc.

Stone, P. J. (2020). Thematic Text Analysis: New Agendas for Analyzing Text Content. In Routledge eBooks. doi:https://doi.org/10.4324/9781003064060-3
Tan, Y. S., Zalzuli, A. D., Ang, J., Ho, H. F., & Tan, C. (2022). Understanding the Workload of Police Investigators: a Human Factors Approach. J Police Crim Psych, 37, 447–456.

Tedesco, D., & Tullis, T. (2006). A comparison of methods for eliciting post-task subjective ratings in usability testing. Usability Professionals Association (UPA), 1-9.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic management journal, 18(7), 509-533.

Tett, G. (2015). The silo effect: The peril of expertise and the promise of breaking down barriers. Simon and Schuster.

Treichler, C. (2019). Consulting Industry and Market Trends: A Two-Sided View, in V. Nissen (ed.), Advances in Consulting Research: Recent Findings and Practical Cases. Cham: Springer International Publishing, pp.253–272. doi:https://doi.org/10.1007/978-3-319-95999-3_12

Tripsas, M., & Gavetti, G. . (2000). Cognition, capabilities and inertia: evidence from digital imaging. Strategic Management Journal, 21(10/11), 1147-1161.

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. Strategic Management Journal, 21, 1147-1161.

Tushman, M., Newman, W., & Romanelli, E. (2004). Convergence and upheaval: Managing the unsteady pace of organizational evolution. Managing Strategic Innovation and Change: A Collection of Readings. New York.

Uotila, J., Maula, M., Keil, T., & Zahra, S. (2009). Exploration, exploitation, and financial performance: analysis of S&P 500 corporations. Strategic Management Journal, 30, 221-231.

Van Teijlingen, E., & Hundley, V. (2002). The importance of pilot studies. Nursing Standard (through 2013), 16(40), 33.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Vennix, J. (1996). Group model building. New York: John Wiley & Sons. Repeated Author. 1999. Group model-building: Trackling messy problems. System Dynamics Review, 15 (4):379-401.

Vennix, J. A., Andersen, D. F., & Richardson, G. P. (1997). Forward: group model building, art, and science. System Dynamics Review, 13 (2):103-106.

Verganti, R. (2009). Design driven innovation: changing the rules of competition by radically innovating what things mean. Harvard Business Press.

Videira, N., Antunes, P., & Santos, R. (2017). Engaging stakeholders in environmental and sustainability decisions with participatory system dynamics modeling. Environmental modeling with stakeholders: Theory, methods, and applications, 241-265.

Videira, N., Antunes, P., Santos, R., & Lopes, R. (2010). A participatory modelling approach to support integrated sustainability assessment processes. Systems Research and Behavioral Science, 27(4), 446-460.

Wadood, F., Akbar, F., & Ullah, I. (2021). THE IMPORTANCE AND ESSENTIAL STEPS OF PILOT TESTING IN MANAGEMENT STUDIES: A QUANTITATIVE SURVEY RESULTS. The Journal of Contemporary Issues in Business and Government, 27(5), 2419–2431.

Walrave, B., Oorschot, v. K., & Romme, A. G. (2010). Ambidexterity and getting trapped in the suppression of exploration : a simulation model. Technische Universiteit Eindhoven, BETA publicatie : working papers; Vol. 314.

Weber, Y. (2019). Mergers and Acquisitions: Exploitation and Exploration. The Cross-Disciplinary Perspectives of Management: Challenges and Opportunities. 29–37. doi:10.1108/978-1-83867-249-220191004

Weick, K. (1995). Sensemaking in Organizations . London: Sage Publications.

Wick, G., & Gampp, F. (2019). System Mapping Toolkit. Retrieved from System Mapping Academy: https://www.system-mapping.com/toolkit
Winter, S. (2000). The satisficing principle in capability learning. Strategic Management Journal, 21, 981-996.

Wolstenholme, E. F. (1999). Qualitative vs quantitative modelling: the evolving balance. Journal of the Operational Research Society, 50, 422-428.

Wright, D., & Meadows, D. H. (2008). Thinking in systems. Earthscan.
XO-Projects. (2024). Customer Journey Mapping Template. Retrieved from https://miro.com/miroverse/customer-journey/.

Zaak, D. N. (2018, februari 8). Accenture - de normaalste zaak. Retrieved from https://www.denormaalstezaak.nl/partners/accenture/#:~:text=Accenture%20Nederland%20telt%20ca%202600,een%20kwetsbare%20positie%20naar%20werk

Zheng, D. L. (2018). Design thinking is ambidextrous. Management Decision, 56(4), 736-756.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007, April). Research through design as a method for interaction design research in HCI. In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 493-502.

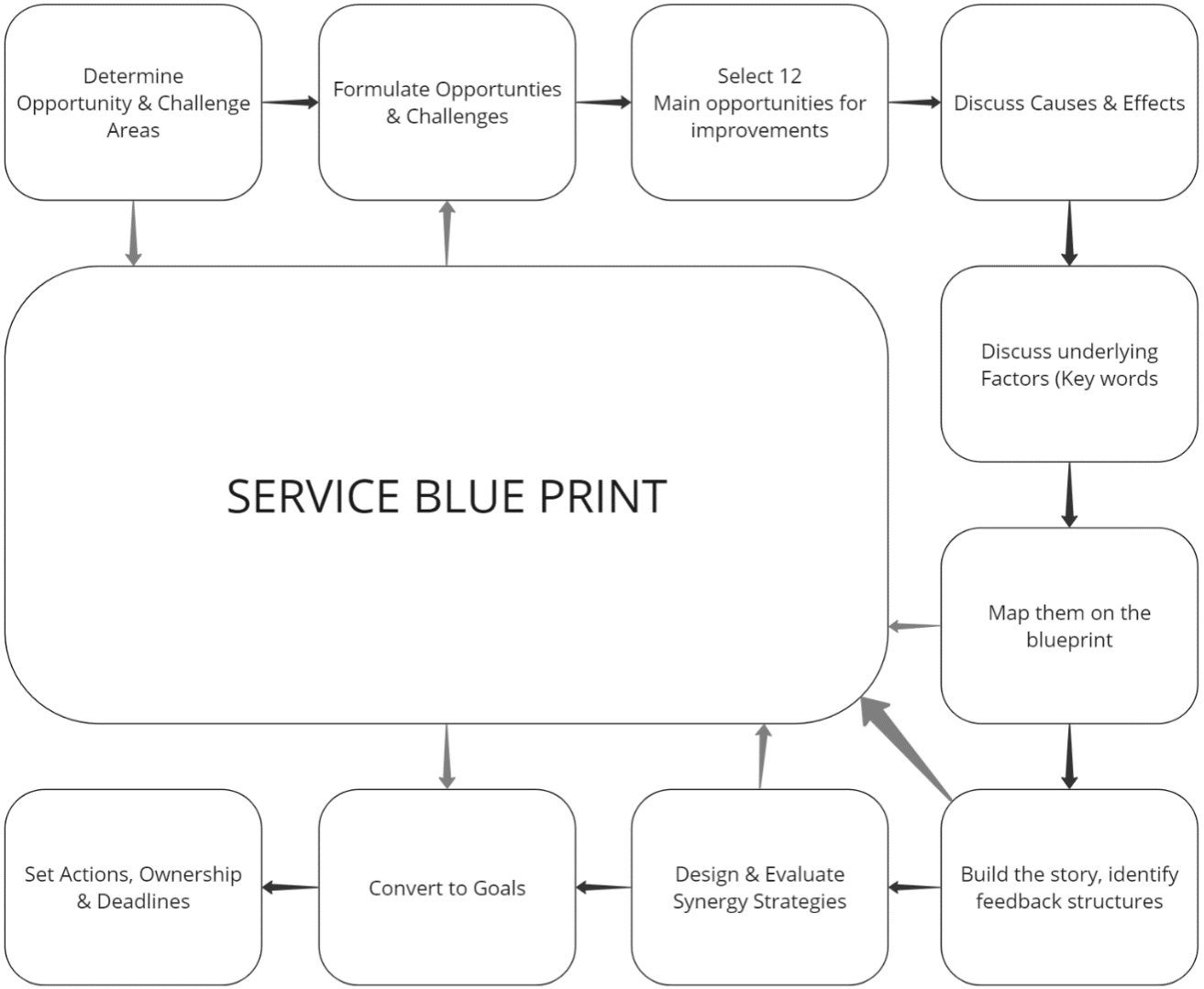
NON-CONFIDENTIAL APPENDIX

Part of: Thesis Co-creating Organizational Synergy by Joris Raaphorst

Contents

1. Design Specifics

1.0 Original Process Structure



1.0.1 Safety protocol

As described in the Ethical Review Form:

The physical risks are comparable to a normal day working at the office. Other risks which must be considered are the potential consequences that participants might face when they openly criticize organizational structure and processes. By anonymously synthesizing the output from the interviews into one service blueprint, it becomes hard to backtrace any specific feedback to a person. Preventing negative association to one particular person.

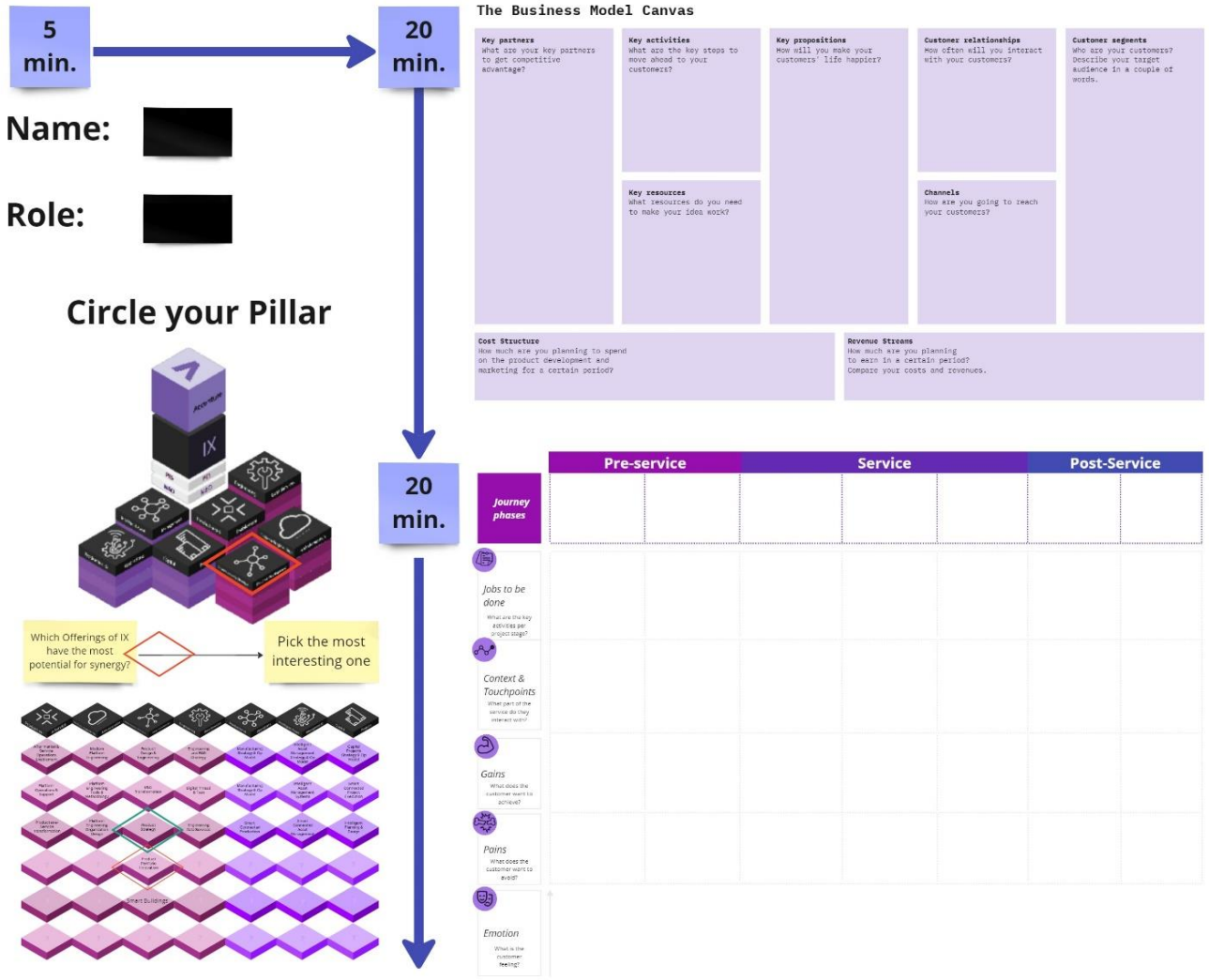
As facilitator it is important to ensure a safe space during the workshop in which no one is pushed to disclose to much, as it might (potentially) put them in a precarious situation. The discussions during the workshop are not recorded. The output of the session is shaped in the form of a causal diagram over who both parties (IX & VB) must agree.

Another risk is that conflict can arise due to organizational differences and underlying tensions. To mitigate this risk the consent form will explicitly mention this risk and underlines the importance of a safe space. It will also refer to a trust person within the organization to facilitate any de-escalation. The importance of a safe space is also mentioned by the researcher during the workshop introduction. No issues are expected, but safe guarding the participants is highly important.

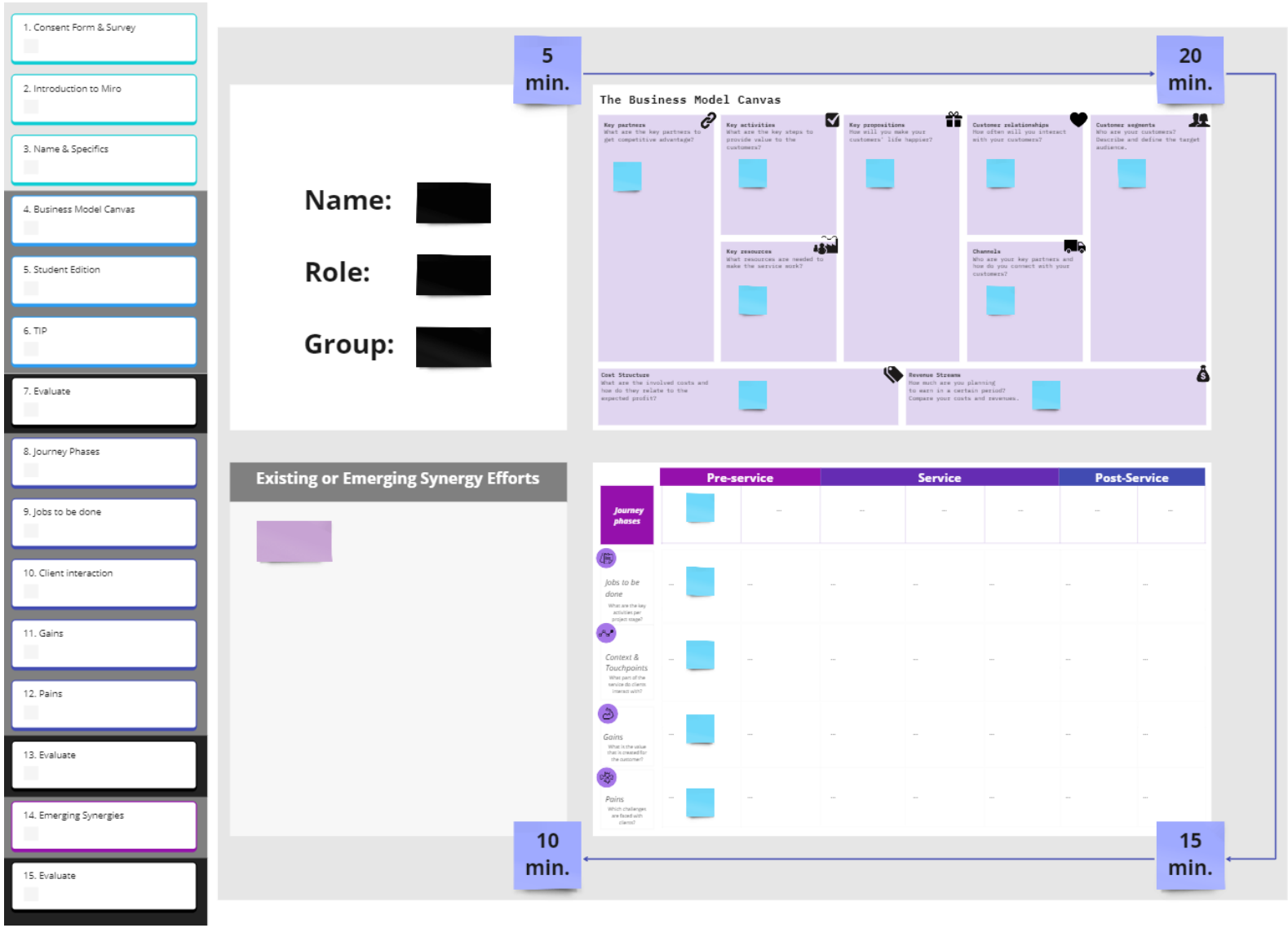
The considerations in the protocol above were integrated in various locations on template as written text to remind participants, as well as it was mentioned in the e-mails and it was explicitly addressed by the facilitator during collaborative sessions.

1.1 Context Capturing Interview

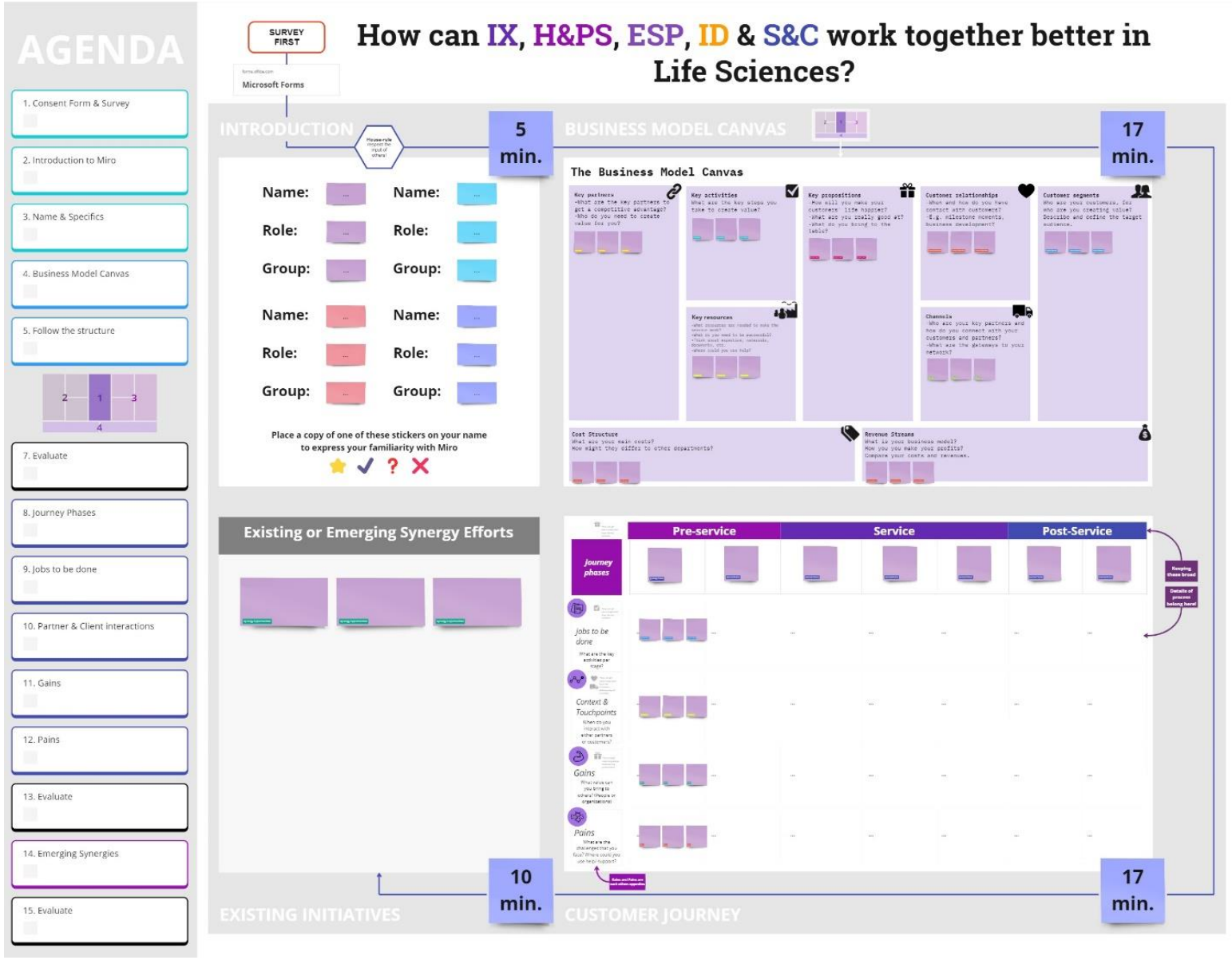
1.1.1 M2.1 Original template design



1.1.2 Pilot template design

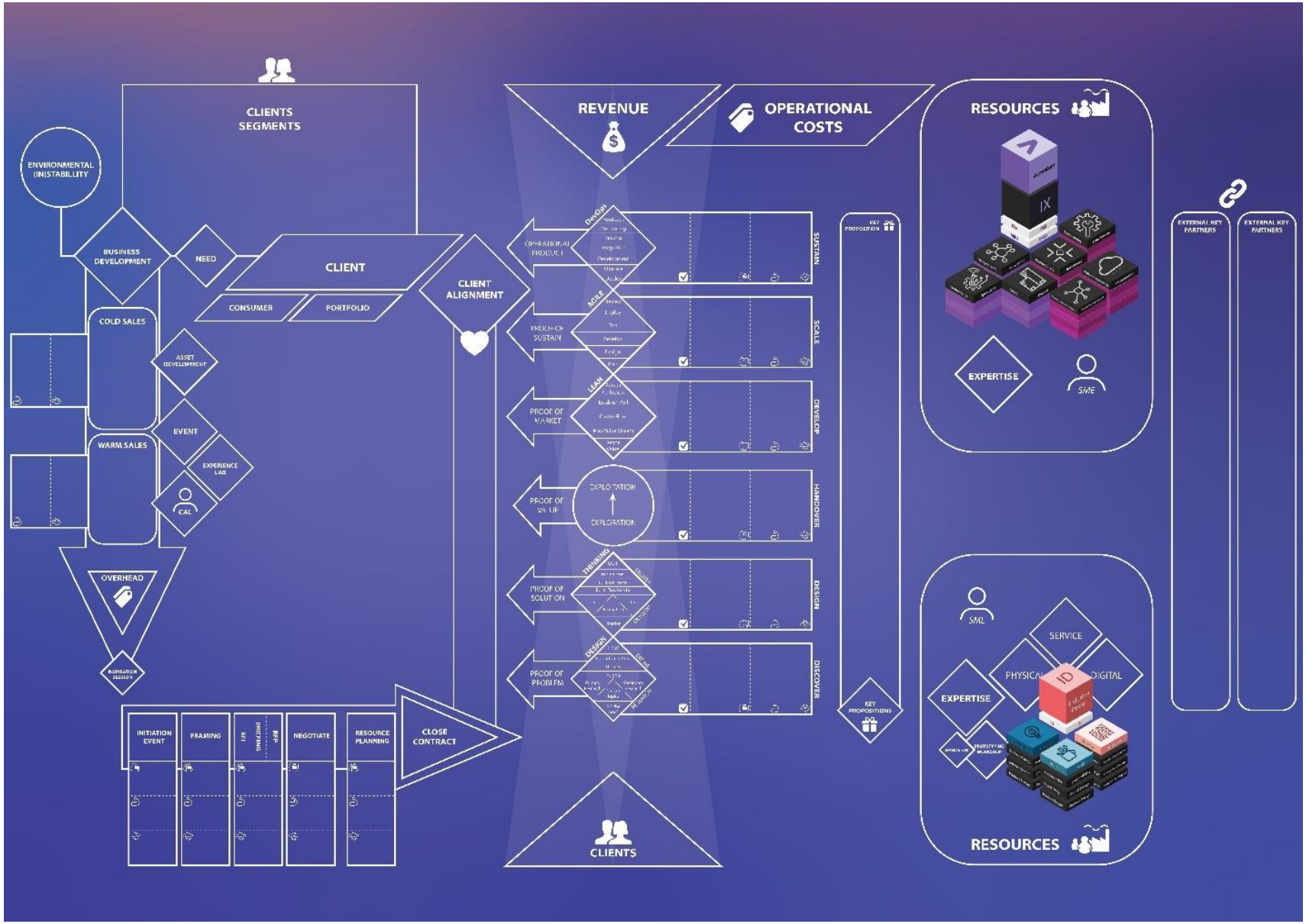


1.1.3 LS – Use-case template



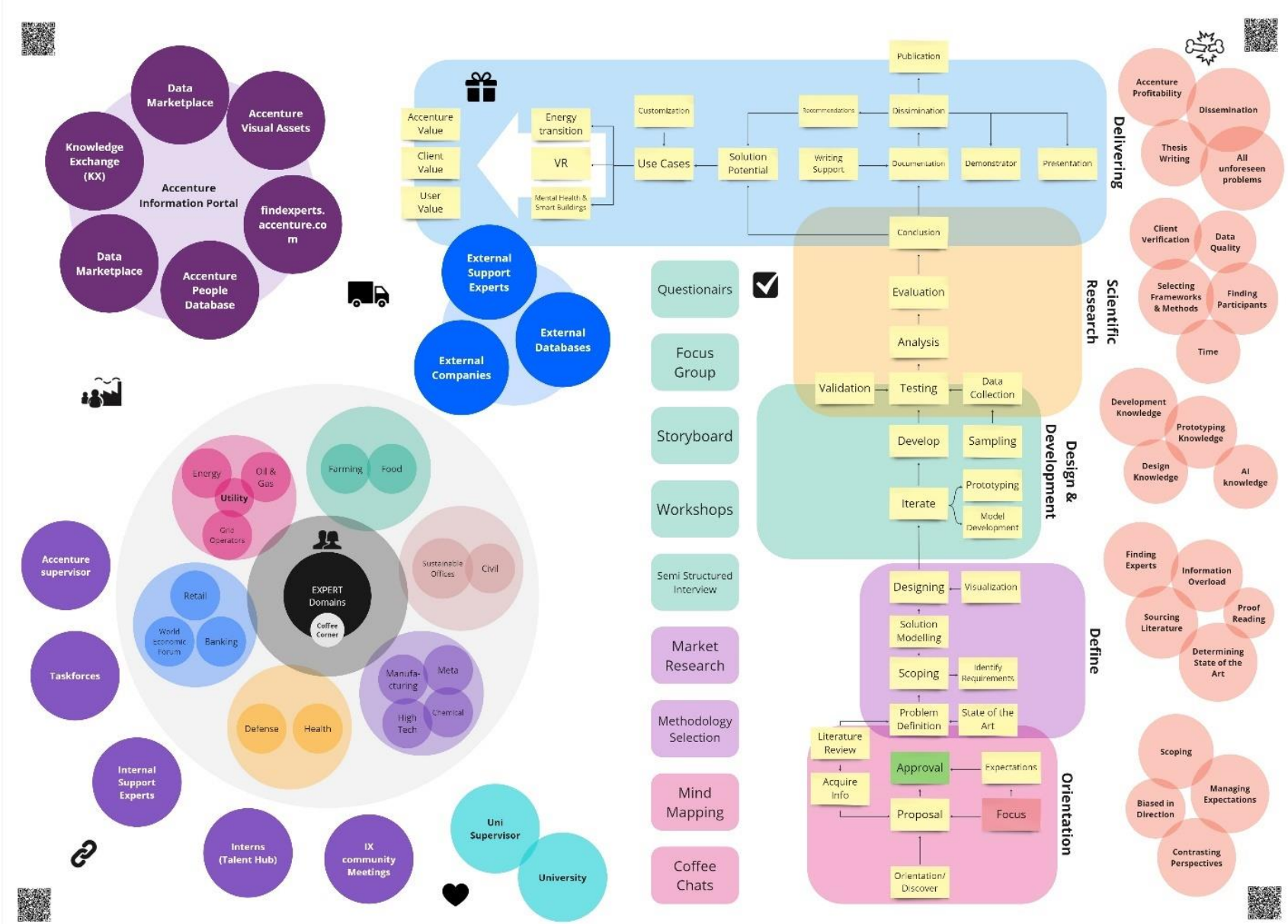
1.2 Blueprint Creation Process

1.2.1 M2.1 Blueprint Design

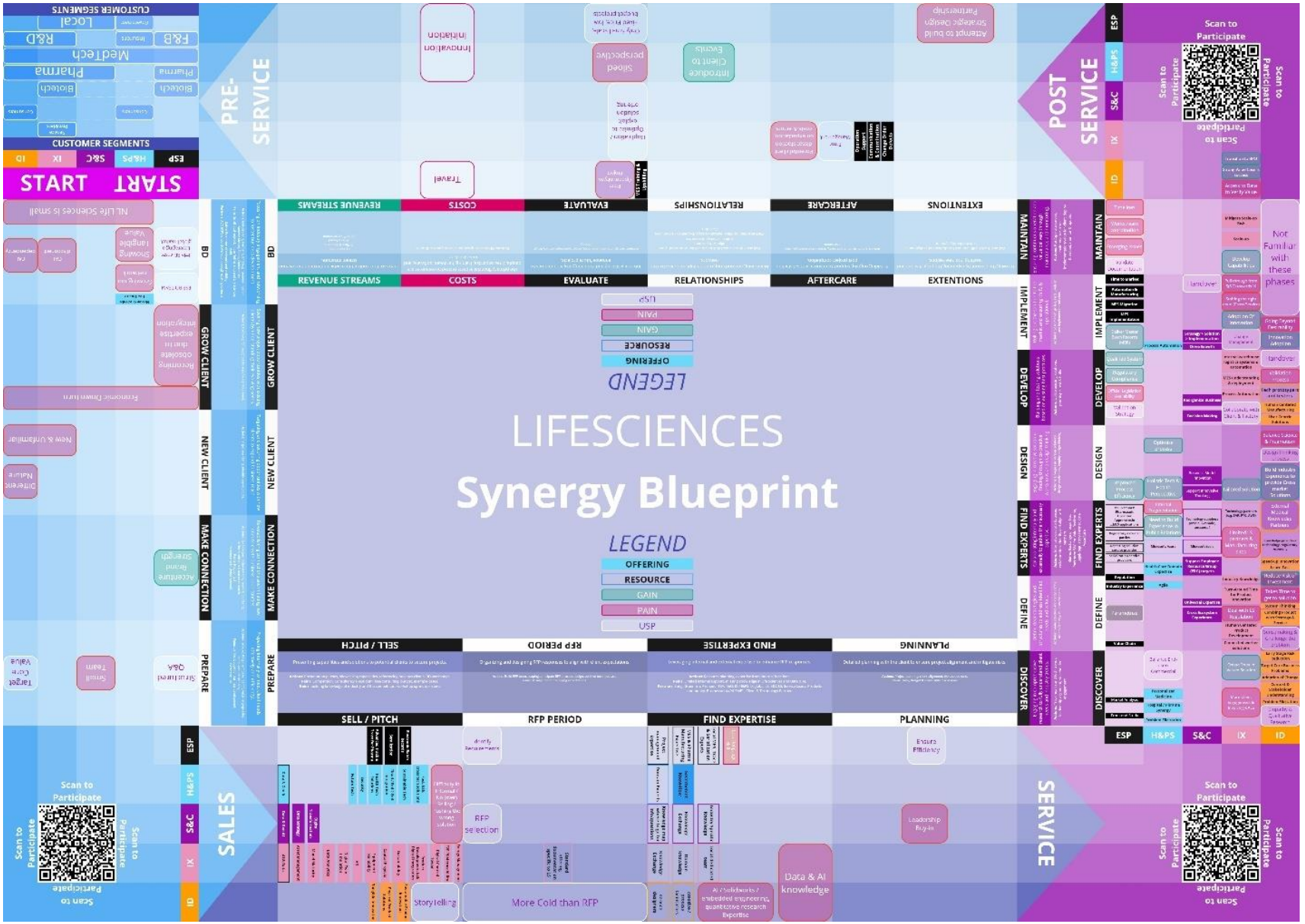


M2.2 – Industrial Design – Final Master Project – Joris Raaphorst – June 2024 – Thesis: Co-creating Organizational Synergy

1.2.2 Pilot Blueprint Design



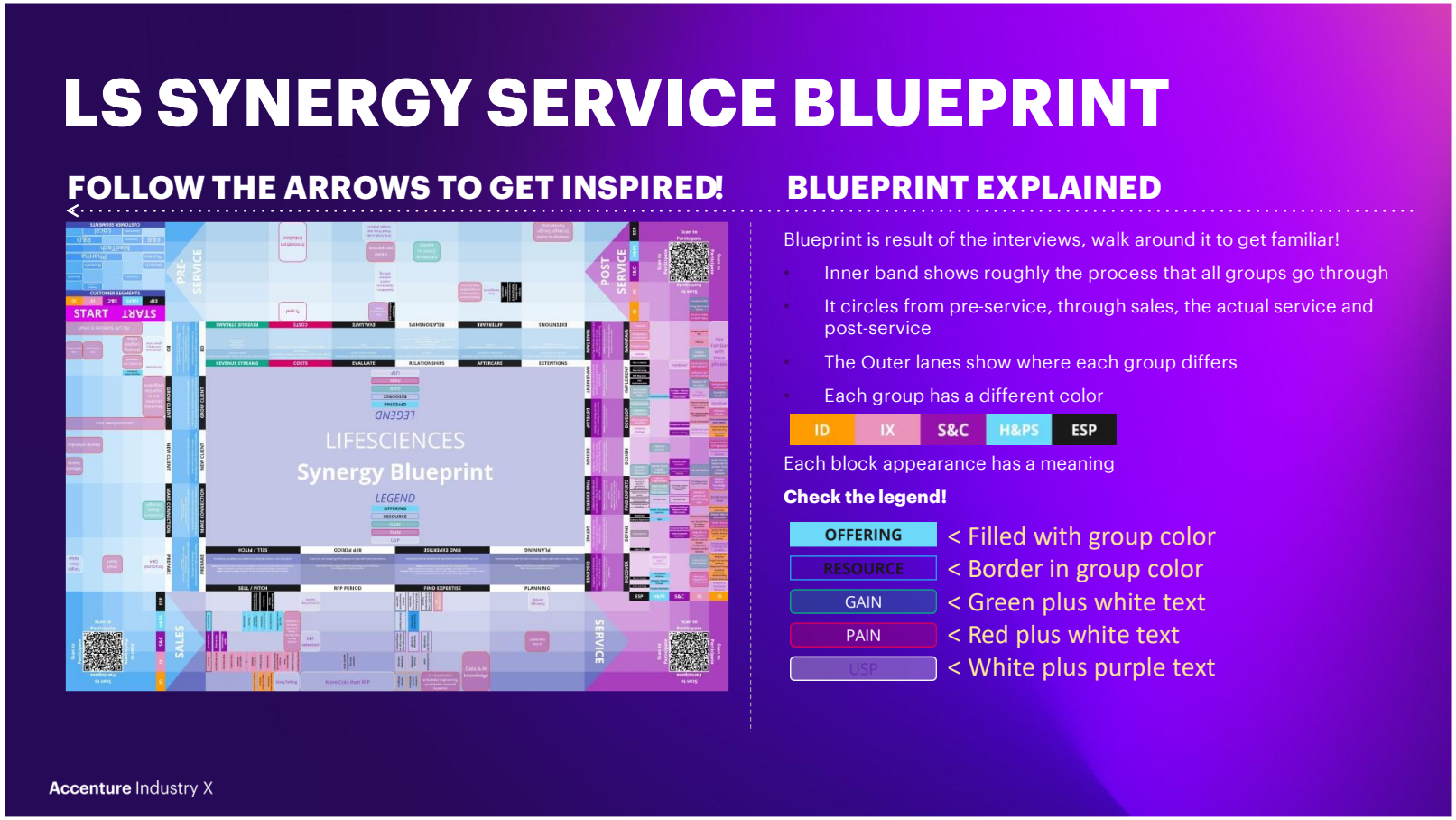
1.2.3 LS - Use-case Blueprint Design



1.3 Opportunity Identification Session

To support incorporating more perspectives, this session will be designed with larger groups in mind that represent a mix of junior and senior participants. The prime focus of this session is to generate a larger quantity of identified opportunities that are properly formulated and prioritized by the group to support the recommended Opportunity Selection after this session by decision-makers.

To ensure a proper user experience that is both feasible and viable, the session will be designed as a one-hour workshop. This will make it easier for participants to attend and reduce the required resources.



1.3.1 Identify Areas of Interest

Firstly, participants are asked to identify areas of interest for collaboration. This can be done by placing an element on the board or circling it with a whiteboard marker. This exercise will allow participants to familiarize themselves with the blueprint and discuss areas of potential collaboration.

1.3.1.1 Facilitator Card 01

01. Identify Area of Interest 15 min.

1. Go through every group, (ID, IX, S&C, H&PS, ESP) ask people to raise their hands if they represent a group.
2. Hand out the black markers
3. Walk around the blueprint, follow the arrows, try to familiarize yourself!
4. Discuss your services.
    1. Ask others about elements on the blueprint.
    2. Try to understand each other's services.

DISCLAIMER: the blueprint isn't perfect and might contain faults, it is intended to spark inspiration and diverge the discussion.

5. Identify areas of interest for synergy, mark them - Use BLACK markers.
6. Collect the black markers.
7. Direct them to the FIRST 4 questions in the survey!

1.3.1.2 Pilot Support Slide

01. Identify Area of Interest

 10 min

Free and open Disucssion

1. Gather around the blueprint that is placed in the middle of the table
2. Discuss your projects, explain elements on the blueprint
3. Identify areas of interest for synergy and mark them on with a marker

Evaluate

1.3.1.3 LS - Use-Case Support Slide

15 min

01. Identify Area of Interest

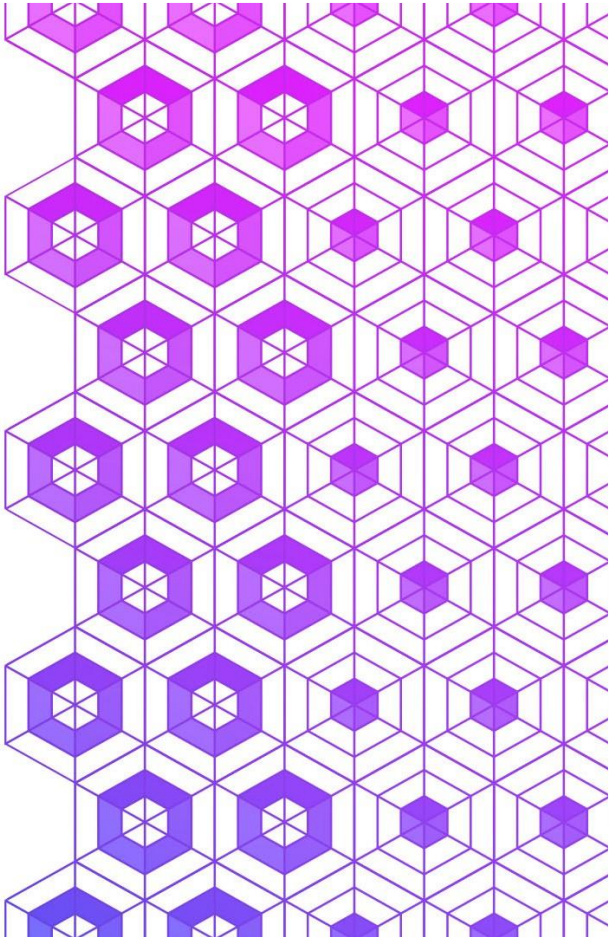
Freely discuss to familiarize

1. Walk around the blueprint, follow the arrows.
2. Discuss your services, ask others about elements on the blueprint.

   DISCLAIMER: the blueprint isn't perfect and might contain faults, it is intended to spark inspiration and diverge the discussion.

3. Identify areas of interest for synergy, mark them - - - Use **BLACK**

Please answer the **first** 4 questions in survey on your phone!



1.3.2 Positioning Emerging Offerings

To improve the quality of the output, the session then continues with positioning existing or emerging strategies on the blueprint. These will be provided in a placeable token on which the offering is briefly formulated. The intention is for the participants to be provided with examples of offerings with high potential and, in the process, familiarize themselves with the blueprint. The new potential might be unlocked by tasking the participants to identify new areas of interest next to the already defined opportunities.

1.3.2.1 Facilitator Card 02

02. Positioning Existing Synergies 10 min.

1. Read out the first existing opportunity, name the involved groups.
2. The group must discuss its position on the blueprint.
3. Include introvert people, all groups should be involved.
4. Wait until consensus is reached.
5. Then place it on the blueprint, or hand it out to someone to place it for you.
6. Direct them to the next 4 questions in the survey!

1.3.2.2 Pilot Support Slide

02. Positioning Emerging Offerings

10 min

Verify existing opportunities

1. Position the existing and emerging synergies on the blueprint.

Evaluate

1.3.2.3 LS - Use-Case Support Slide

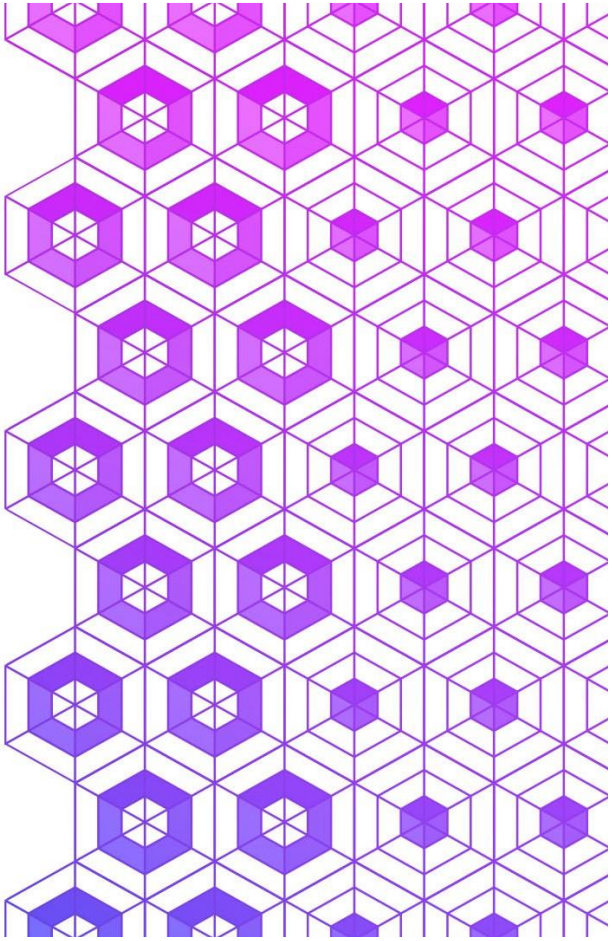
10 min

02. Positioning Existing Synergies

Become aware of existing synergies

1. The facilitator will read out an existing synergy.
2. Discuss as group its position on the blueprint
3. Facilitator places it on the blueprint

Please answer the **next** 4 questions in survey on your phone!



1.3.3 Identify New Opportunities

After positioning the existing or emerging offerings, participants are provided with empty tokens, like those containing emerging offerings. Participants are tasked to formulate new opportunities not covered by the emerging offerings. Utilizing more participants to formulate and merge opportunities can result in more nuanced opportunity descriptions incorporating various perspectives.

1.3.3.1 Facilitator Card 03

03. Identify New Opportunities 15 min.

1. Place the hexagons on the table, and hand out the purple markers.
2. Hypothesize / speculate about potential, new synergy/ collaboration opportunities.
3. Prompt the group to start with considering areas of interest, but everything is allowed.
4. They have to write them down on a hexagon – Use PURPLE!
5. And position them on the blueprint.

TIP: Is the group getting stuck/ stagnating? Support teaming up with someone, or group discussion, talking helps to go further!

6. Collect the purple markers.
7. Direct them to the next 4 questions in the survey!

1.3.3.2 Pilot Support Slide

03. Identify New Opportunities

10 min

Expand Opportunity Profile

1. Now formulate new synergy opportunities and place them on the board

Evaluate

1.3.3.3 LS – Use-case Support Slide

15 min

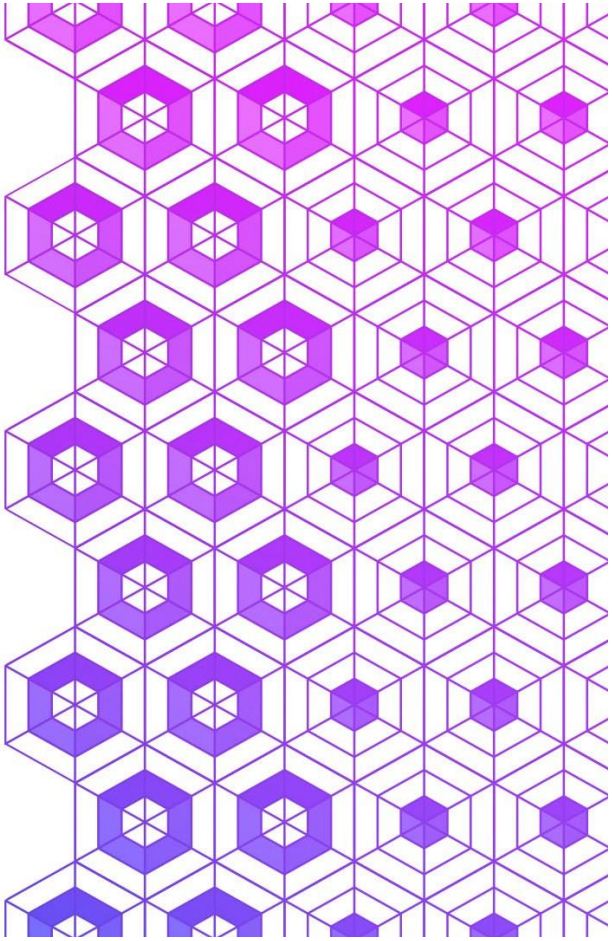
03. Identify New Opportunities

Verify existing opportunities

1. Consider (open) areas of interest
2. Hypothesize new synergy opportunities
3. Write them down on a hexagon – Use **PURPLE**!
4. Position them on the board

TIP: Are you getting stuck? team up with someone, talking can help!

Next 4 questions in survey on your phone!



1.3.4 Pre-Selection

Lastly, participants are provided with thirty sticker dots in three colors, ten of each. Each color represents either desirable, feasible, or viable. They are asked to distribute the dots across the opportunities and, by doing so, perform a hands-on evaluation of the opportunities on these factors. As such, a prioritization of the opportunities can be made in which these three factors are considered. Prioritizing the opportunities makes it easier and more valuable to present this list of opportunities to decision-makers.

1.3.4.1 Facilitator Support Slide 04

04. Selection **10 min.**

1. Hand out the red and blue markers.
2. Clearly explain that they need to place one dot per opportunity!
3. **5 DESIRABILLTY DOTS - USE RED MARKER**
4. **5 FEASIBILLITY DOTS - USE BLUE MARKER**
5. Collect the markers when participants are finished to keep track of who is finished.
6. Direct them to the next 4 questions in the survey!

1.3.4.2 Pilot Support Slide

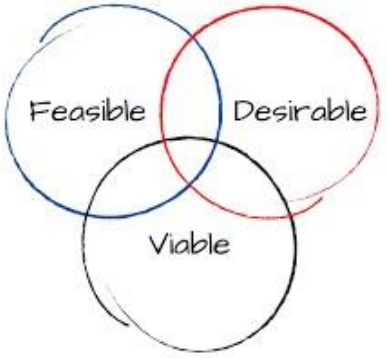
04. Selection

10 min

Power Dotting rounds on Desirabillity & Feasibility.

1. When ready, distribute 5 dots to feasible opportunities
2. Lastly, do the same with 5 dots to viable opportunities

Evaluate



1.3.4.3 LS – Use-case Support Slide

10 min

04. Selection

Power Dotting rounds on Desirabillity & Feasibility.

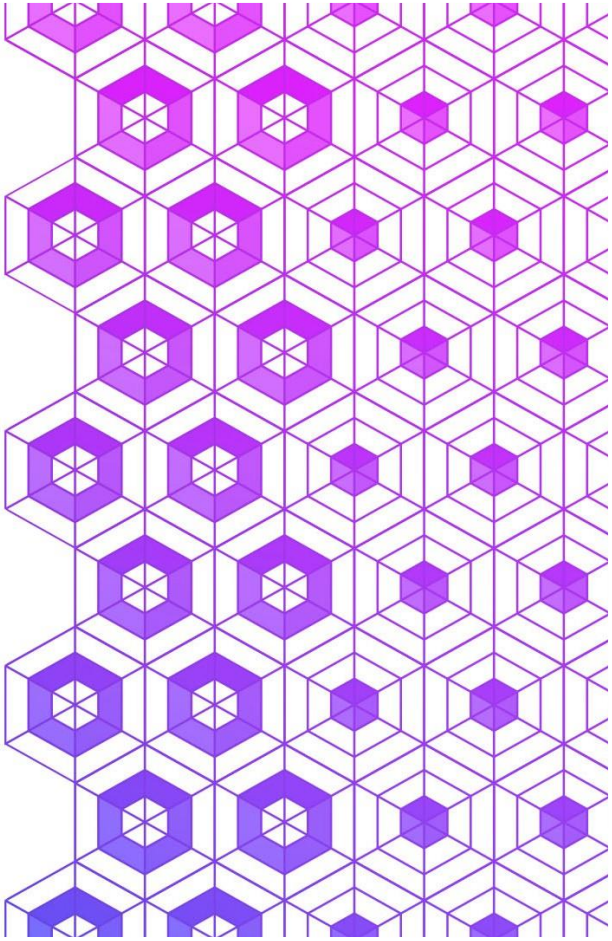
1. Per person distribute:

●●●●● **5 DESIRABILLTY DOTS** **- USE RED MARKER**

●●●●● **5 FEASIBILLITY DOTS** **- USE BLUE MARKER**

Only use one dot per opportunity!

Next 4 questions in survey on your phone!



1.3.5 Opportunity Framing

To prepare the opportunities for the next phase, the group is tasked to form an alliance of 2-3 people based on the identified opportunities by placing a token with their alliance name on it to block it.

Then, this alliance is given a 'Framing template' on which they can elaborate on their opportunity. This Framing template is based on the starting template for the Participatory Modelling method used in the modelling phase.

1.3.5.1 Facilitator card 05

05. Opportunity Framing End – 5 min.

7. Point out a few high-ranking opportunities and explain that we are going to select and frame some opportunities that are worth exploring.
8. This is the time to form alliances.
    1. Explain that they carefully must choose an opportunity they believe should be explored with one other participant from another group!

(INITIATE ONLY GROUPS OF TWO!)

    2. Ideally, only groups of 2 are formed, but since some opportunities require a multistakeholder perspective, it is okay to form larger groups, but only with permission from the facilitator!

9. Point towards the purple markers, tokens pens & templates.
10. Explain that they have to grab a token, name their opportunity & place it down.
11. After they can grab a template and a pencil to frame the opportunity.
12. If they have time, they may frame another opportunity!
13. Ask them to finalize the rest of the survey, this takes about 6 minutes.

1.3.5.2 Pilot support slide

05. Opportunity Framing

 10 min

1. Chose an opportunity you want to pursue and link up with a person that represents the other perspective.
2. Together frame the opportunity using the provided template, you have five minutes.
3. Now select another opportunity and repeat the process.

Evaluate

1.3.5.3 LS – Use case support slide

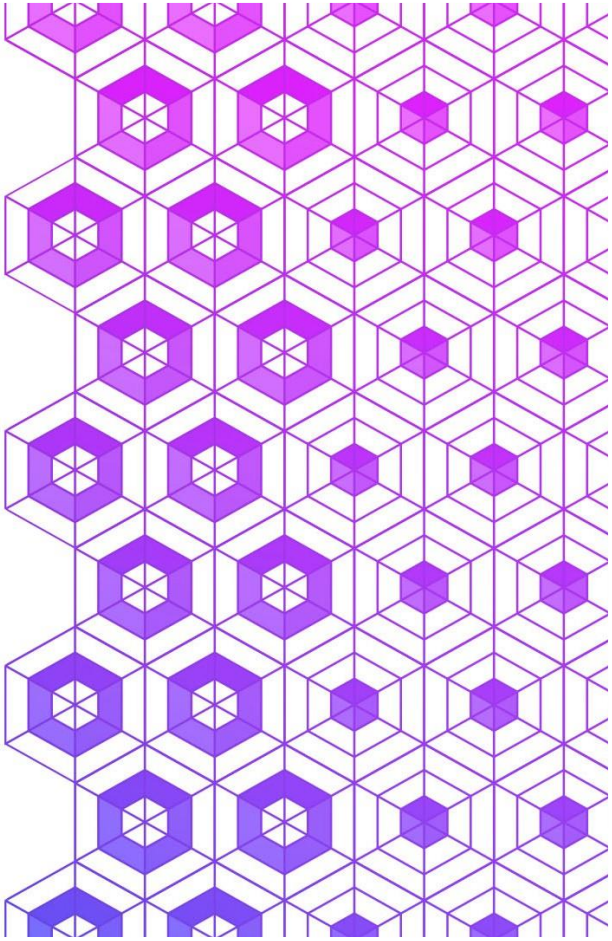
15 min (finish at 4:25)

05. Opportunity Framing

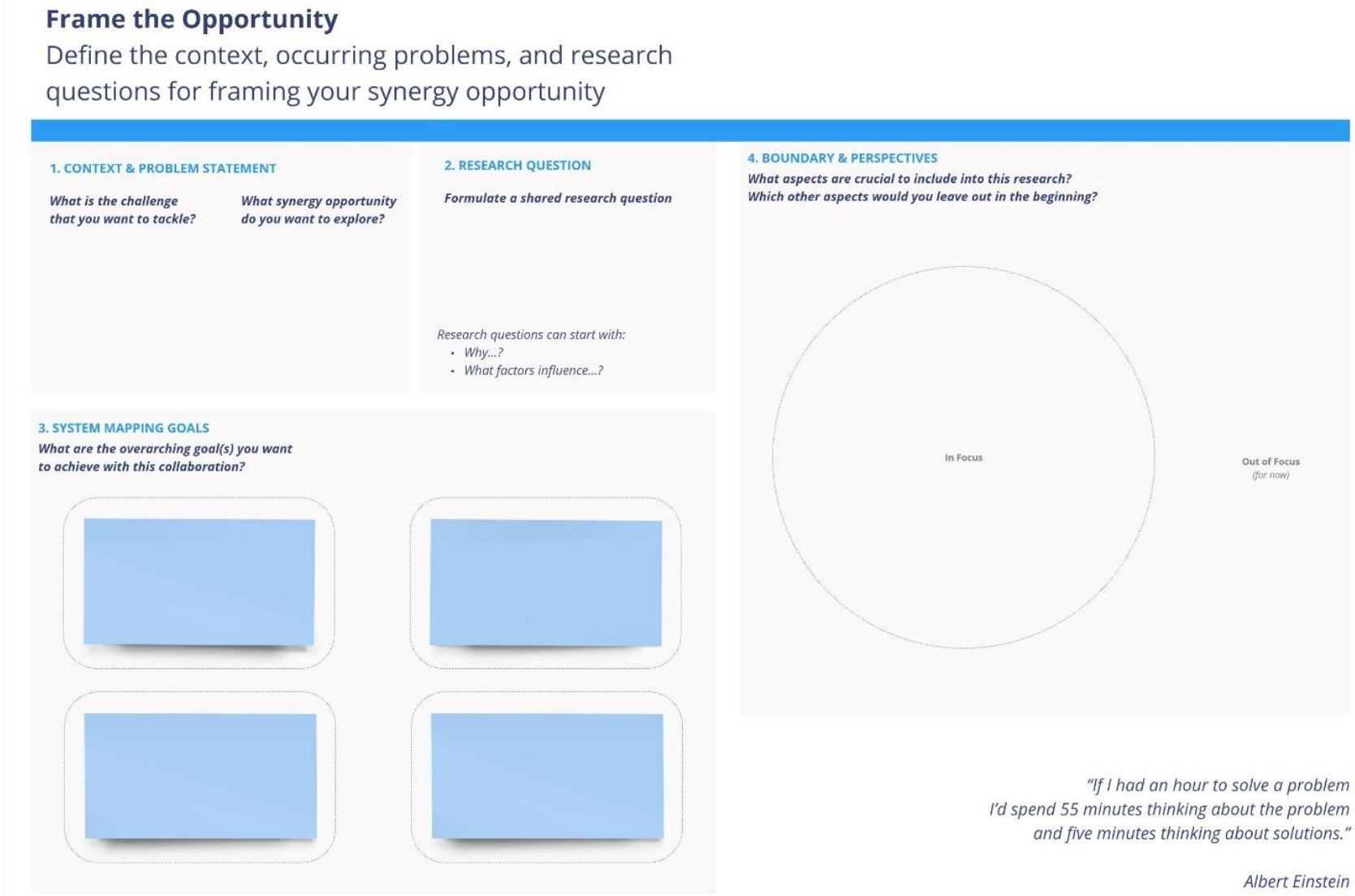
ONLY WORK IN PAIRS FROM DIFFERENT GROUPS

1. Carefully chose an opportunity, link up with another that represents a different perspective.
2. Grab a token, name your opportunity & place it.
3. Together frame the opportunity using the provided template
4. If you have time, frame another opportunity!

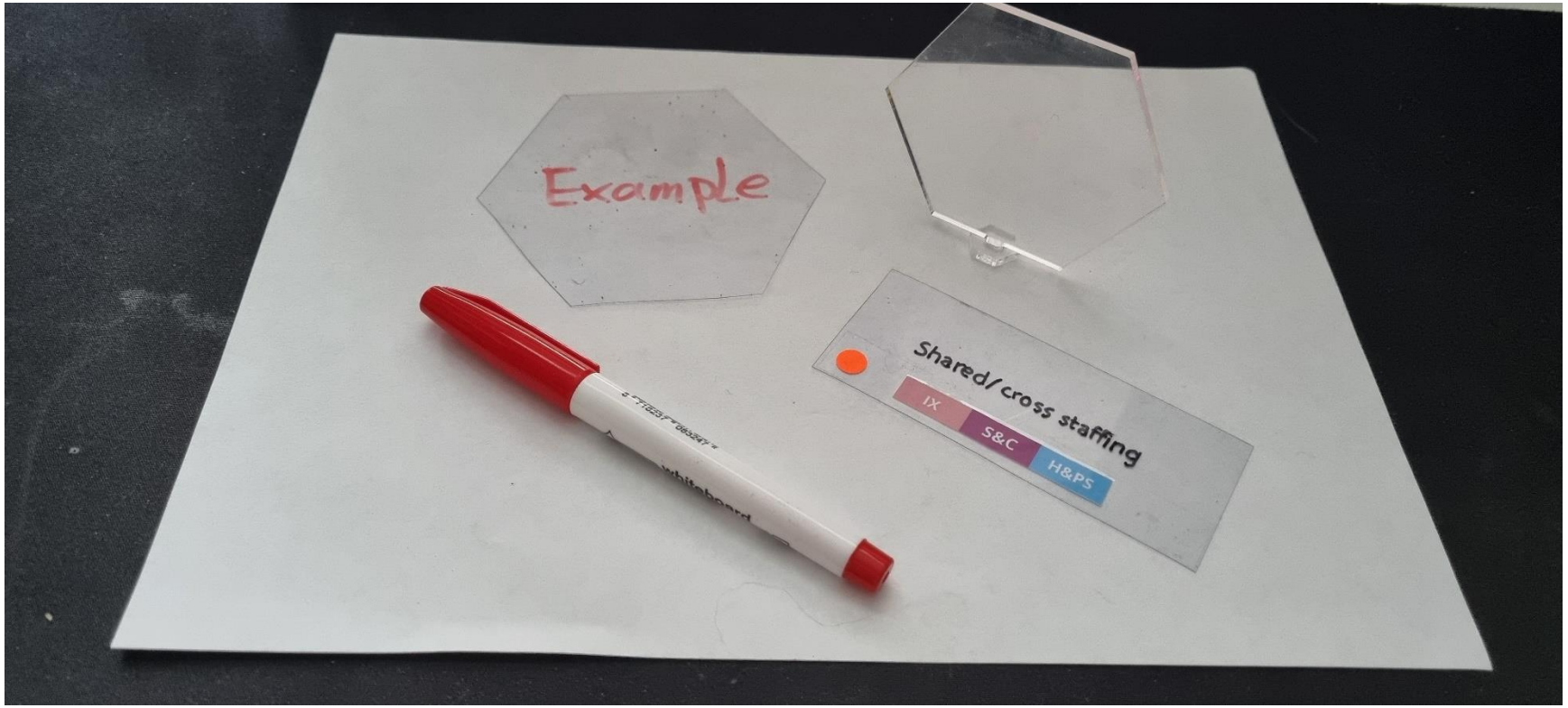
Finish the survey on your phone!



1.3.5.4 Opportunity Framing Template



1.3.6 Physical Tools

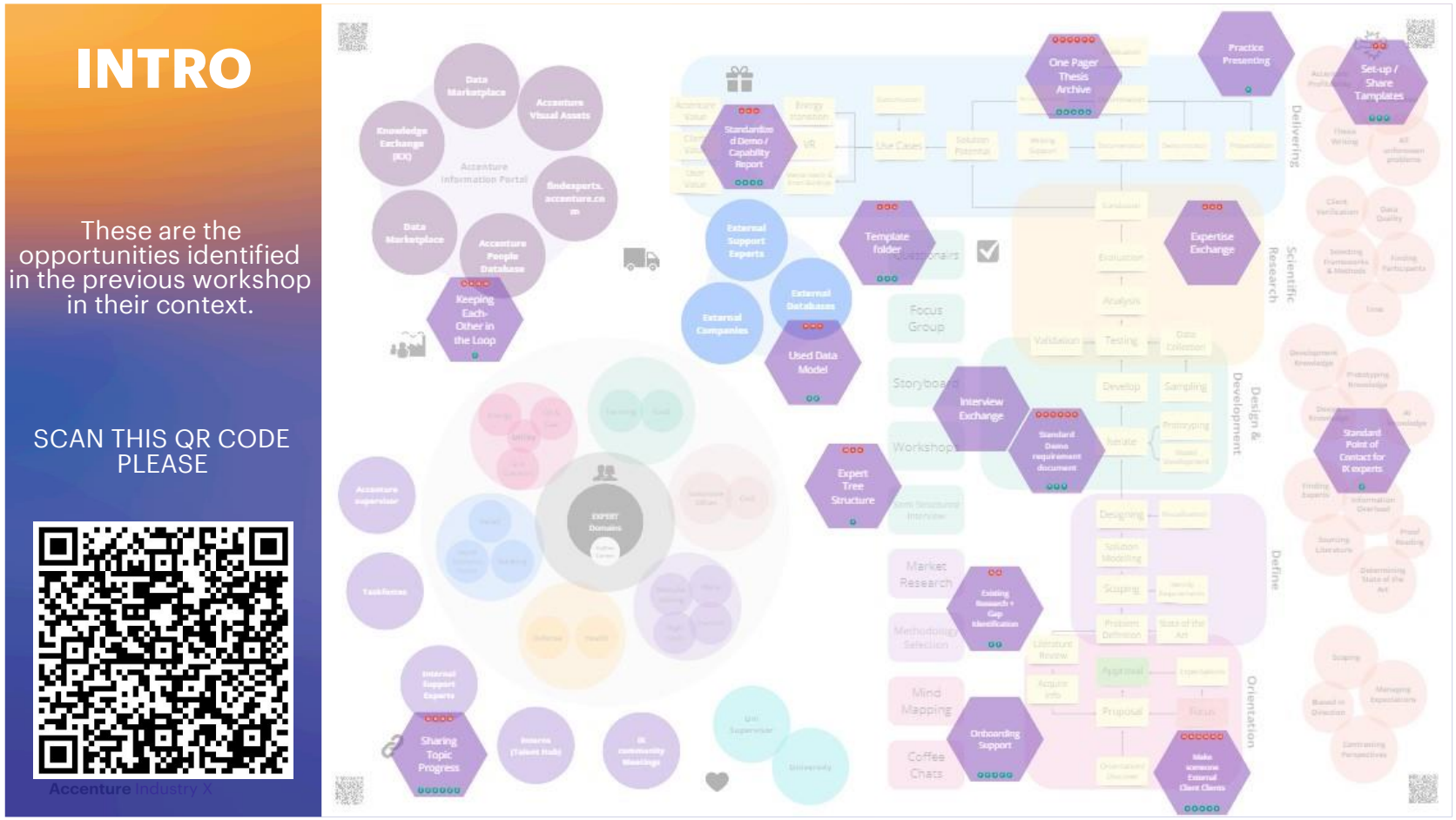



POST
SERVICE

1.4 Opportunity Selection

1.4.1 Introduction

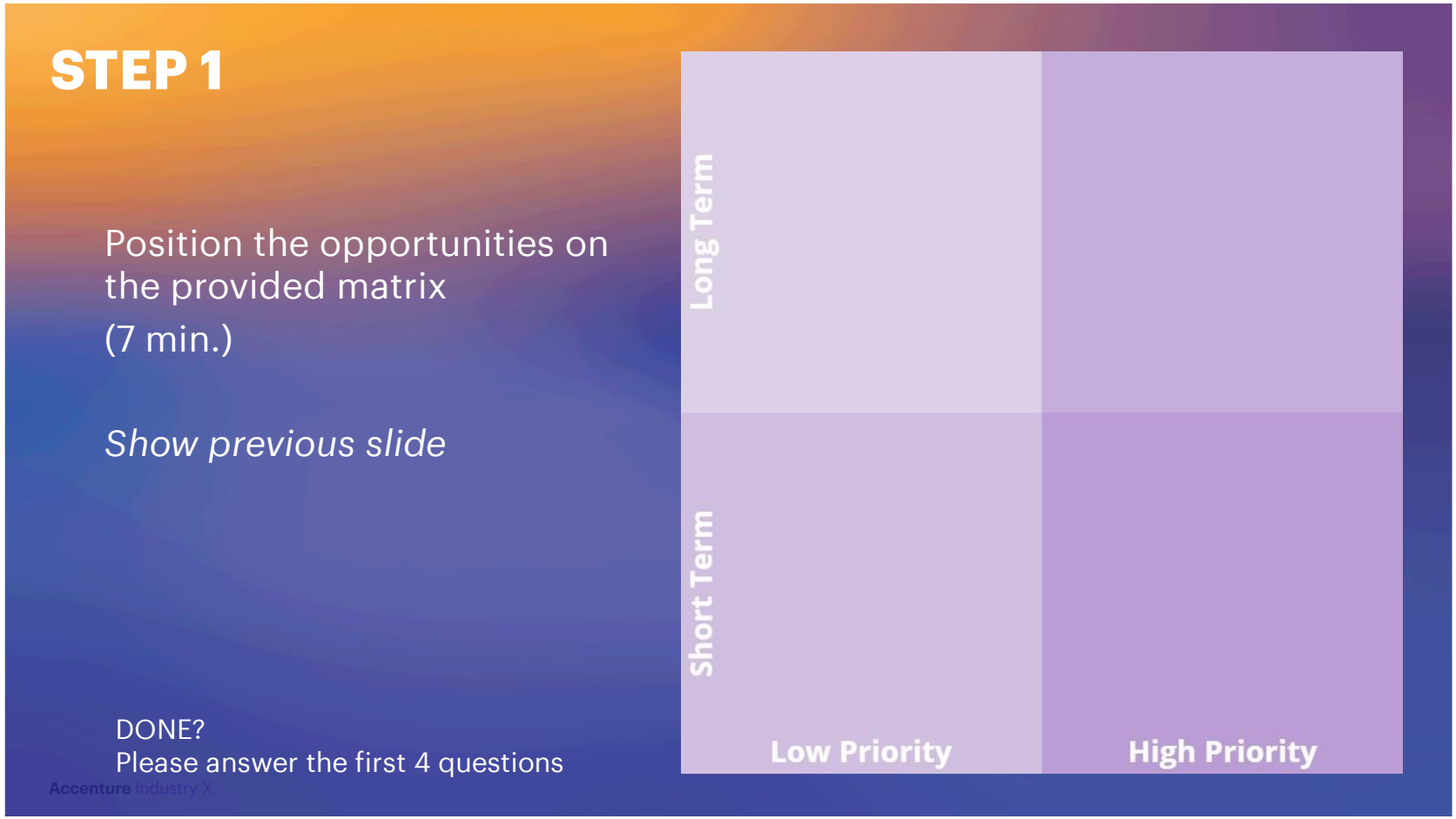
1.4.1.1 Pilot Introduction



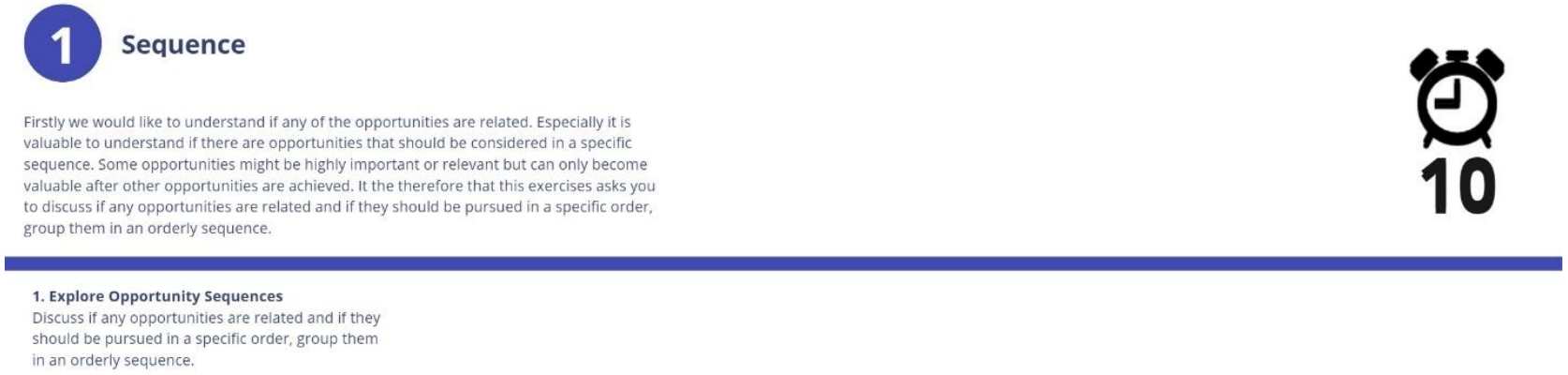
1.4.1.2 LS-Use Case Miro template Introduction

1.4.2 Step 1 – Grouping & Matrix Mapping

1.4.2.1 Pilot Explainer – Step 1 - Mapping

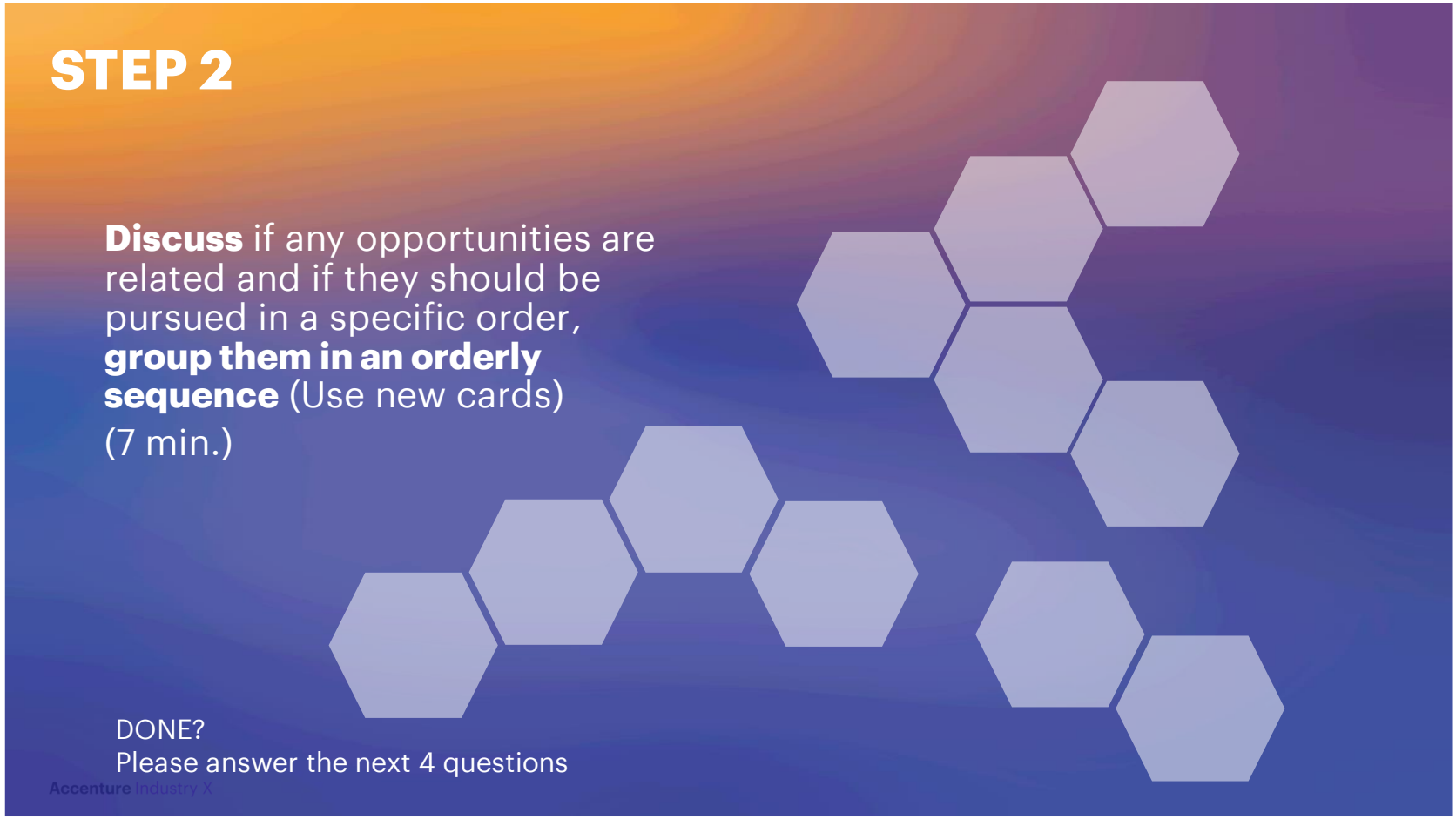


1.4.2.2 LS-Use Case Miro template Step 1



1.4.3 Step 2 – Sequencing & Matrix Mapping

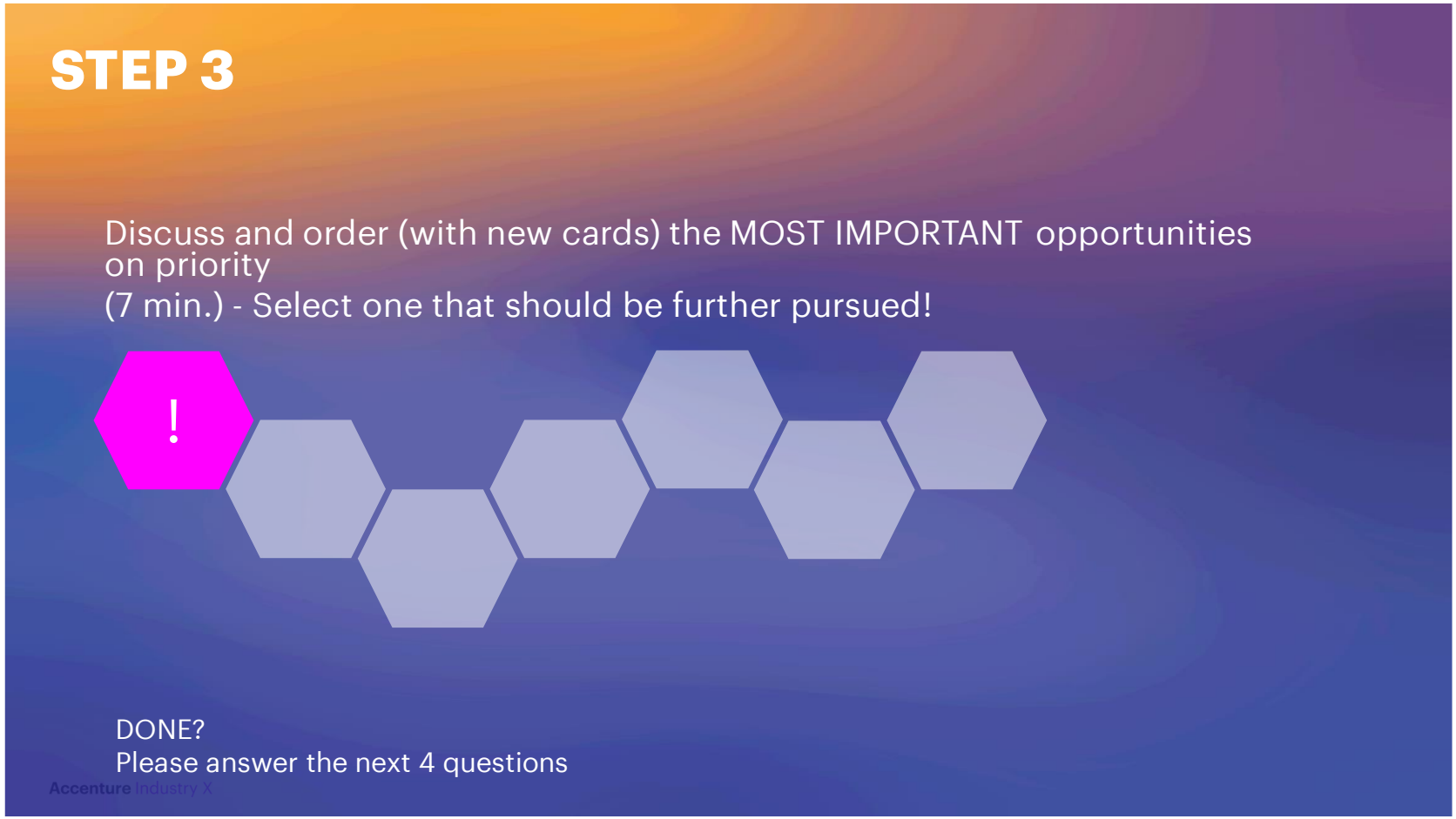
1.4.3.1 Pilot Explainer – Step 2



1.4.3.2 LS-Use Case Miro template Step 2

1.4.4 Step 3 – Prioritization list selection

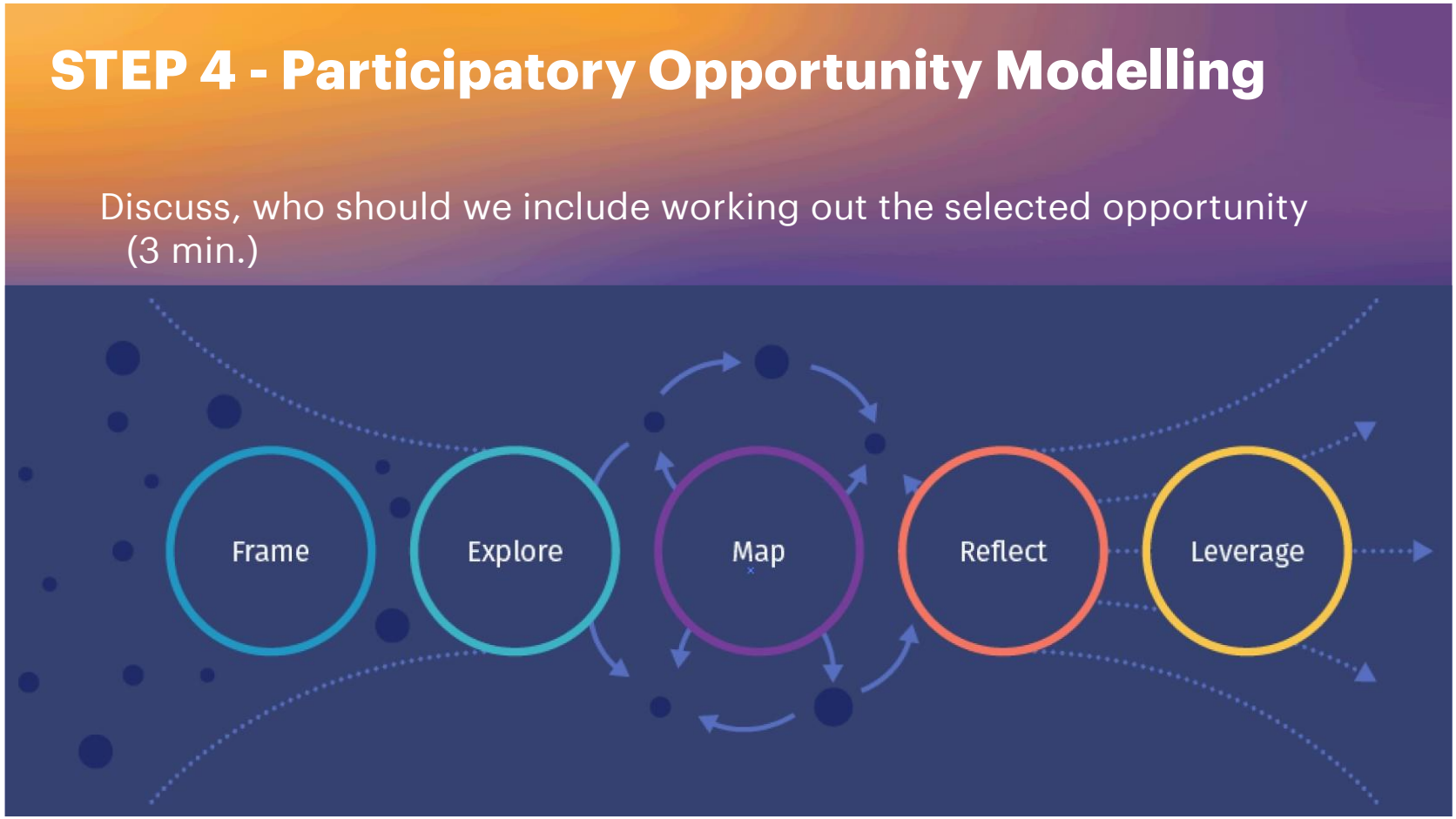
1.4.4. Pilot Explainer – Step 3



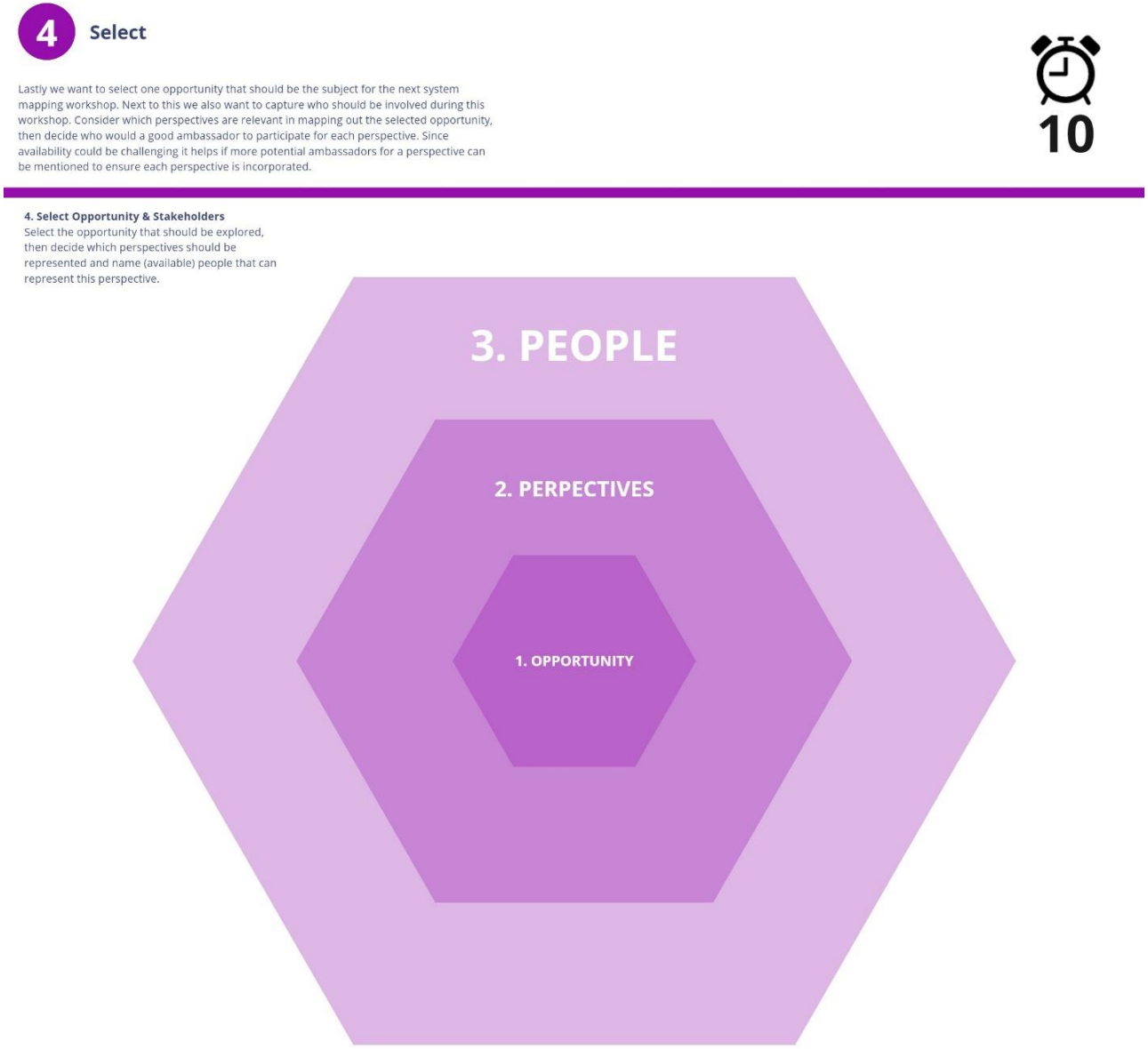
1.4.4.2 LS-Use Case Miro template - Step 3

1.4.5 Step 4 - Recruitment

1.4.5.1 Pilot Explainer – Step 4



1.4.5.2 LS-Use Case Miro template - Step 4



1.5 Opportunity Modelling

For this session, the toolkit from https://www.system-mapping.com/toolkit was adopted one-on-one. Please check out this toolkit on its creator's website. To get an impression of the template, please open the attached censored template from the Use-Case session.

[Right Click → Object → Open]

Censored
LS-OM-Session Temp

2. Data Collection – Survey example

This document includes the data collection survey and Consent Form. To review them, please open the Example-Data collation template.

[Right Click → Object → Open]

Example-Data
collection template - S

3. Statistical Analysis

3.1 R-Studio Code

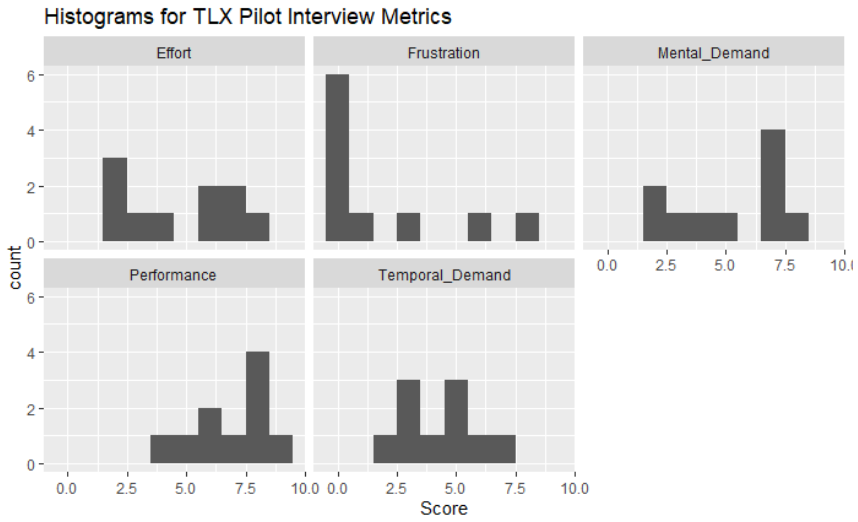
The R-code and its output were rendered in the following document for those who wish to review my approach to coding the statistical analysis. The statistical analysis's output is also extracted from this analysis to provide a clearer and more structured overview.

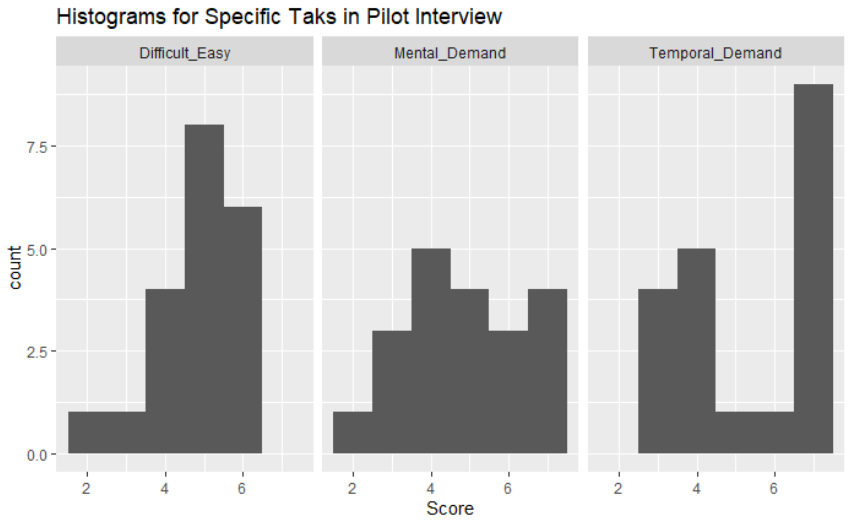
[Right Click → Object → Open]

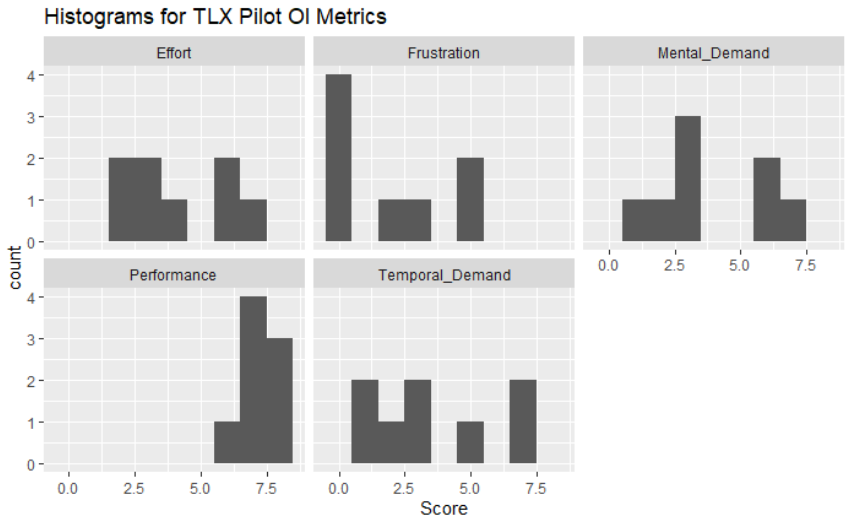
Data_Analysis_Maste
r.pdf

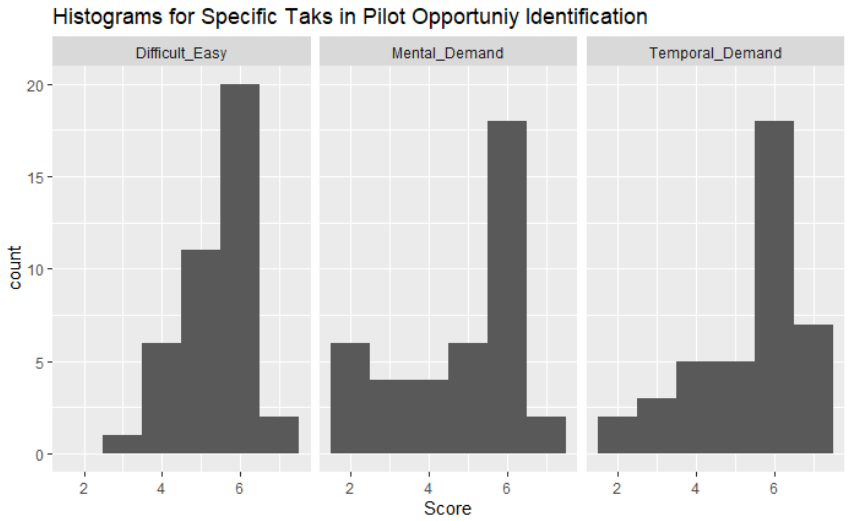
3.2 Histograms – Data distribution check

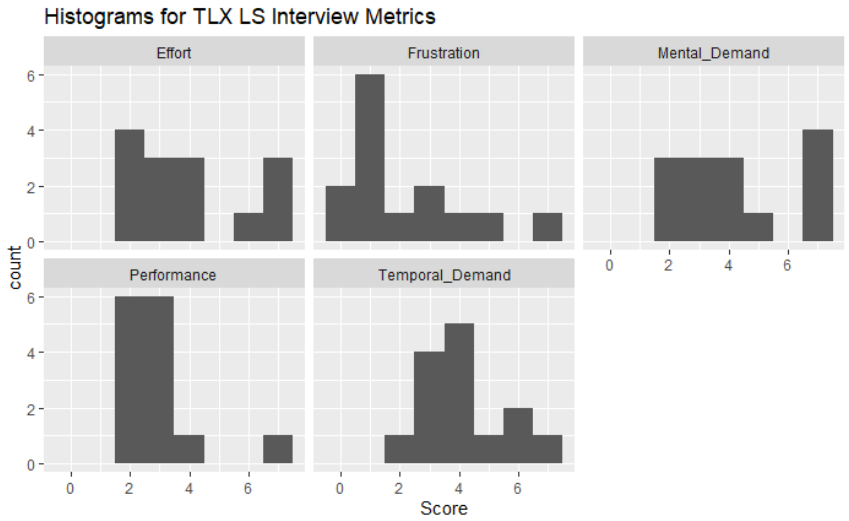
To check if the data was normally distributed, histograms were generated for the Task-Specific and Task Load data of the Pilots and Life Science use-case interviews and opportunity identification. As no normal distribution was identified, this study assumes a non-normal distribution for every session.

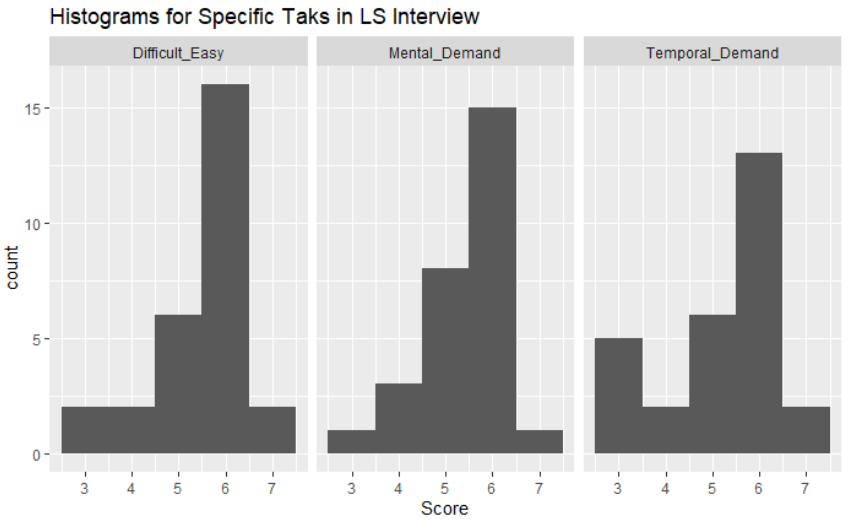


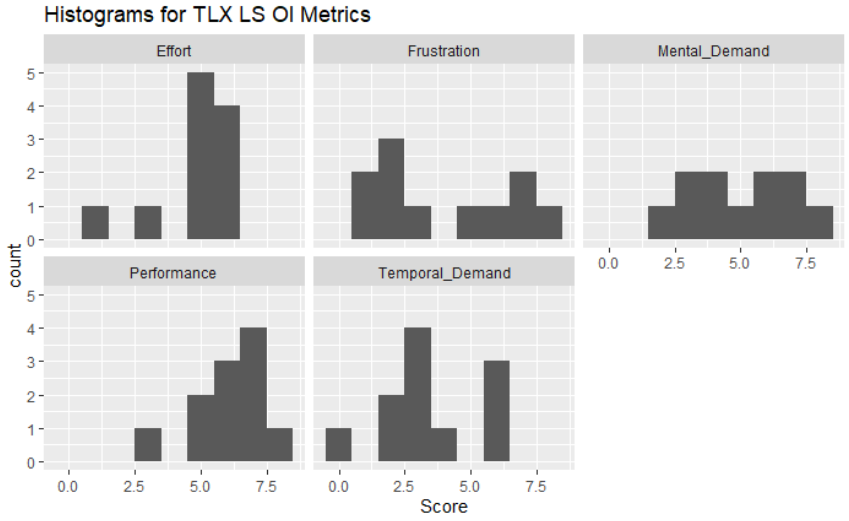


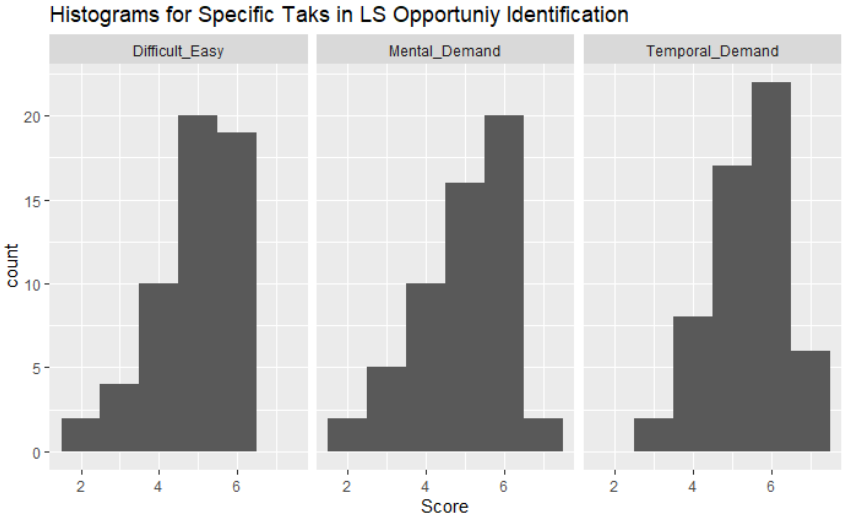





Histograms for TLX LS Interview Metrics
Effort
Frustration
Mental_Demand
Performance
Temporal_Demand
count
Score

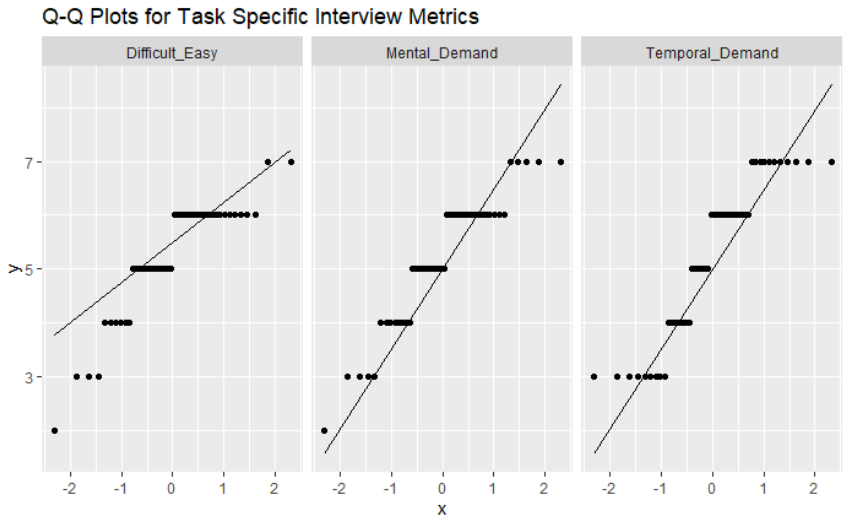

Histograms for Specific Taks in LS Interview
Difficult_Easy
Mental_Demand
Temporal_Demand
count
Score

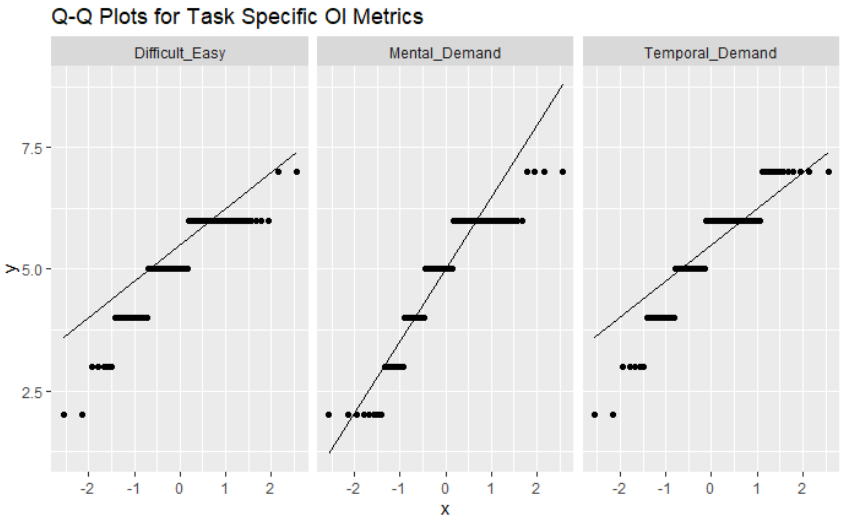

Histograms for TLX LS OI Metrics
Effort
Frustration
Mental_Demand
Performance
Temporal_Demand
count
Score

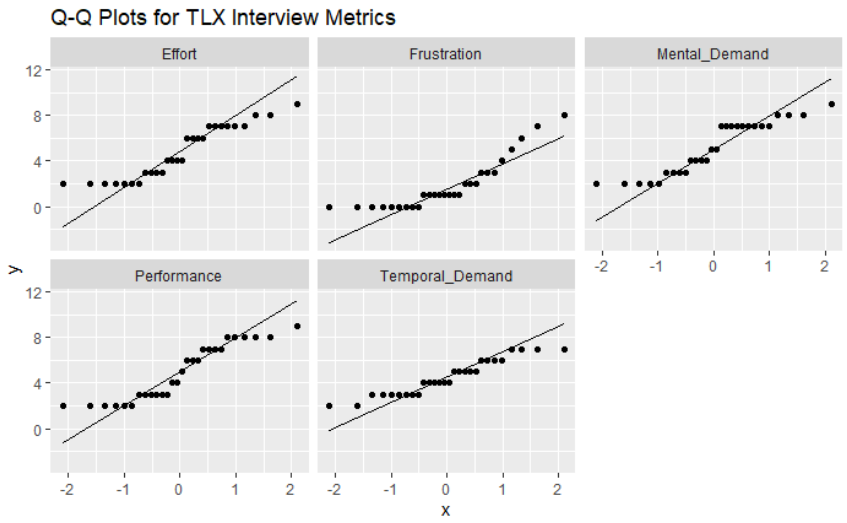

Histograms for Specific Taks in LS Opportuniy Identification
Difficult_Easy
Mental_Demand
Temporal_Demand
count
Score

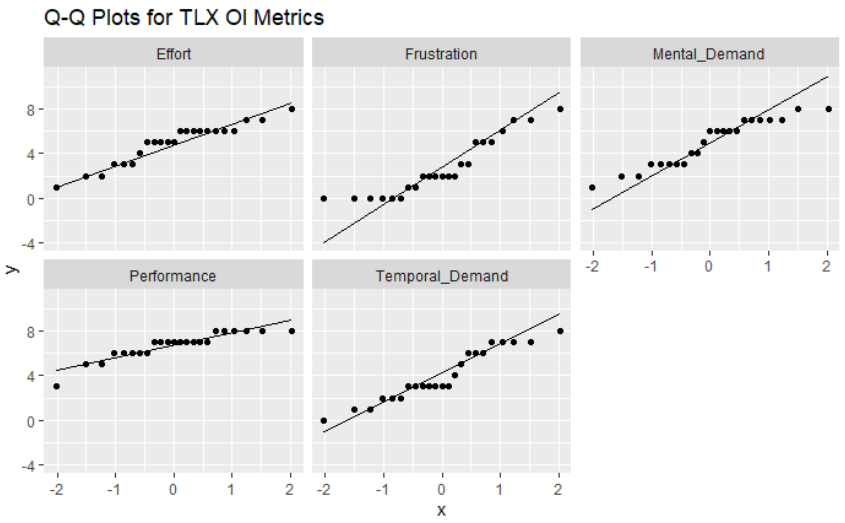
3.3 Q-Q plots – Normality assessment

To check if data was normally distributed, a Q-Q plot was generated for the Task-Specific and Task-Load data of the Pilots' interviews and opportunity identification. As no normal distribution was identified, this study assumes a non-normal distribution for every session.









4. Pilot Data Analysis

The purpose of this analysis was to use the collected data to identify design improvements to the workshop sessions and the research protocol. Additionally, it helps to acquire a general understanding of the workshop's performance in this specific use case.

4.1 Research Protocol

During the pilot, each participant was asked post-survey, in a semi-structured interview format, how they experienced the survey and if they had any remarks regarding its length, difficulty, and clarity. These questions and answers were audio-recorded and transcribed.

The feedback of the pilot participants on the survey was remarkably positive, as some participants put it:

(1) "It definitely is one of the better forms that I have seen, and better surveys done by interns or from others for that matter".

(2) "I think the time indication is good for all the Miro board elements and also the survey."

(3) "It's a fine form. Not too much writing, but more like decisions and that's better I guess."

(4) "No, it was not difficult. It was quite easy."

(5) "it was straightforward and clear."

(6) "I think it's very intuitive and easy to fill in."

In summary, **participants judged the survey as good, fast, easy, intuitive, and with a decent number of questions**.

Participants, however, also expressed constructive feedback, based on which I improved the survey. Firstly, a clarification will be included in the survey to notify participants about the difference in scale direction, as some participants mentioned that they initially assumed that all questions were structured in one direction (e.g., left = negative, right is positive). This is not surprising as the TLX scales are equal in direction, but the direction of the UEQ scales is randomized.

"The positive number was high, and then the other questions were the other way around. And also, with the questions about the two opposites. Like, nice or not nice. Sometimes the negative one was on the left side and the positive one was on the right side, and sometimes the other right.

When I had the last question of the first page, I thought I did it wrong because I didn't, checked enough if the positive one was on the left or the right side. But I did do it right but, so that was something I noticed that it's always changing and everything is very low and very high."

Secondly, a participant mentioned skipping the reading of the explanation text in the survey. Making parts of the explanation bold could support its readability, allowing participants to scan over it without missing the essence of the text.

"I think I may have skipped over a little bit of the introduction reading part, so if there was something really important then I might have skipped that. I don't personally feel like that's a problem. It might be a problem for you as a researcher if people don't completely read. It, but I think it's"

Other feedback involved topics that couldn't be changed, as they are inherently part of the standardized survey, and changing them would disrupt the comparability with the available benchmark data. These will be brought to attention in the discussion segments as research limitations. For instance, the meaning of words was sometimes unclear or not perfect, which might cause deviations in the survey.

"where it asked about 'vertrouwen' and the English version was secure or not secure, but did it mean like trustworthiness or? "

Or that the translation of some words from the UEQ didn't always feel correct.

"Translations weren't always correct. I feel like. For example, there was one that said new, which was translated to inventive. It doesn't feel right to me. I would more say novel and new, for example, where inventive is more, it's more than new, it's doing. It's not only new but doing something with the new in my vocabulary so. Can be a bias on my end. But something to have a look at, maybe. And as always, sometimes the scales just don't feel intuitive."

Lastly, it became clear that the participants' context might influence the measures. A participant's situation might be more complex or evolved than others, influencing its difficulty or task load. This difference should be considered in the interpretation of the results.

4.2 Structured Interview

The structured interviews were performed by interns (60 min.) and their supervisors (45 min.). The design change implemented was the incorporation of a modified Delphi-inspired method, tasking supervisors to review the interns' data input. When disregarding the effects of any contextual differences between the benchmark from Raaphorst (2023), the interns' results should be similar. However, a difference between the interns and Supervisors is expected as the nature of their exercises has changed.

The first hypothesis is, therefore, that:
The interns in the pilot study will experience a similar task load as the participants in the Raaphorst (2023) benchmark study.
The second hypothesis is that:
The supervisors in the pilot study will experience a lower task load as the participants in the Raaphorst (2023) benchmark study.
The third Hypothesis is then that:
The supervisors will experience a lower task load compared to the interns in the pilot study.

4.2.1 End of Survey Comments

4.2.1.1 Staring Participants (Interns)

Do you have any tips?	Do you have any tops?	Other feedback?
No	Makes you think about the process of delivering value	
It would be easier to do a bit further in the process, but I guess some people are much further in their topic	I like visual thinking in miro, it can be very useful also later in the process especially if you also see the miro of other interns	
No	I think the timing was really good!	
Having more time to think about it is good, maybe tell interns to think about their research and what surrounds it (without telling them about the study if it could reduce validity?)	It was well organized and quite smooth. The Miro board looked great!	
Some parts had some repetition, their descriptions could be made more distinctive.	Super clear set up, good guiding descriptions under the labels, and it flowed well from one activity to the next.	
Be clear from what perspective you write the map/journey. More synergy actions	Good Guidance, clear miro board	

4.2.1.2 Follow-up Participants (Supervisors)

Do you have any tips?	Do you have any tops?	Other feedback?
- put the agenda of the session in the Miro board - make sure the post-its are ready for the people to use in the correct color	Clearly explained and there was enough time	
Maybe some examples beforehand - For some items, such as customer value etc, you first need to get "rolling"	Good approach, usefull, creative, well worked out and prepared. Good time indication, no unrelevant questions.	
Check the scales and translations	Very well prepared, good facilitation. All in all a well done exercise.	no
Update the part where it talkes about the client for the pains to what you told me when I asked you about it :)	Good moderation on Joris' side, Very clear and visually pleasing in structure of the Miro board.	

4.2.2 NASA Task Load Index

4.2.2.1 P-TLX – Interviews – Interns vs Bench

No significant difference is found when the task load of interns, supervisors, and the Raaphorst (2023) benchmark is compared. As expected, a clear difference is visible between the experienced task load of the supervisors and that of the benchmark. This difference is primarily found in the required effort, but temporal and mental demand also strongly differ and are near the significance threshold.

This difference is also visible between interns and supervisors, but to a lesser extent, as the interns reported a slightly lesser task load than the benchmark. This slight difference can have many reasons (e.g., the reduced complexity of the interns' projects compared to the context of the Raaphorst (2023) participants, reduced intrinsic motivation, reduced will to perform, or familiarity with the template or Miro).

Based on these results, we can conclude that the change in the exercise significantly reduced the required effort and mental and temporal demands of the reviewing participants **without reducing the perceived performance of the participant**. These results are achieved regardless of reducing the duration of the supervisor exercise to 15 minutes, making this a more feasible approach for deployment with higher-level professionals for whom time is scarcer. It can, however, not be stated that the results are equally valuable, as the perceived performance does not reflect this. Which was also recognized by one of the supervisors *"I personally would say, do 2 blank inputs so me not seeing [Interns] responses and then just thinking of my own and then, post, you combine the two, would be maybe a better approach for having novel input".*

TLX Metric	Mean Intern	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	5,7	7,5	-1,8	6,0	0,232	0,500
Performance	7,0	7,0	0,0	12,0	1,000	0,000
Mental Demand	5,8	8,0	-2,2	3,5	0,075	0,708
Temporal Demand	4,7	6,3	-1,6	5,0	0,157	0,583
Frustration	2,0	1,3	0,8	11,0	0,912	0,083

4.2.2.2 P-TLX – Interviews – Supervisors vs Bench

TLX Metric	Mean Supervisor	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	3,3	7,5	-4,3	0,5	0,041	0,938
Performance	6,8	7,0	-0,3	9,0	0,879	-0,125
Mental Demand	4,3	8,0	-3,8	0,5	0,041	0,938
Temporal Demand	3,8	6,3	-2,5	1,0	0,053	0,875
Frustration	1,5	1,3	0,3	5,5	0,536	0,313

The significant difference between the supervisors and the benchmark in the effort, mental demand, and temporal demand indicates that the newly implemented modified Delphi method positively impacts the task load of the follow-up participants.

4.2.2.3 P-TLX – Interviews – Interns vs Supervisors

TLX Metric	Mean Interns	Mean Supervisors	Mean Difference	W	P-Value	Effect Size
Effort	5,7	3,3	2,4	19,5	0,128	-0,625
Performance	7,0	6,8	0,3	13,0	0,912	-0,083
Mental Demand	5,8	4,3	1,6	17,0	0,321	-0,417
Temporal Demand	4,7	3,8	0,9	16,0	0,444	-0,333
Frustration	2,0	1,5	0,5	14,5	0,631	-0,208

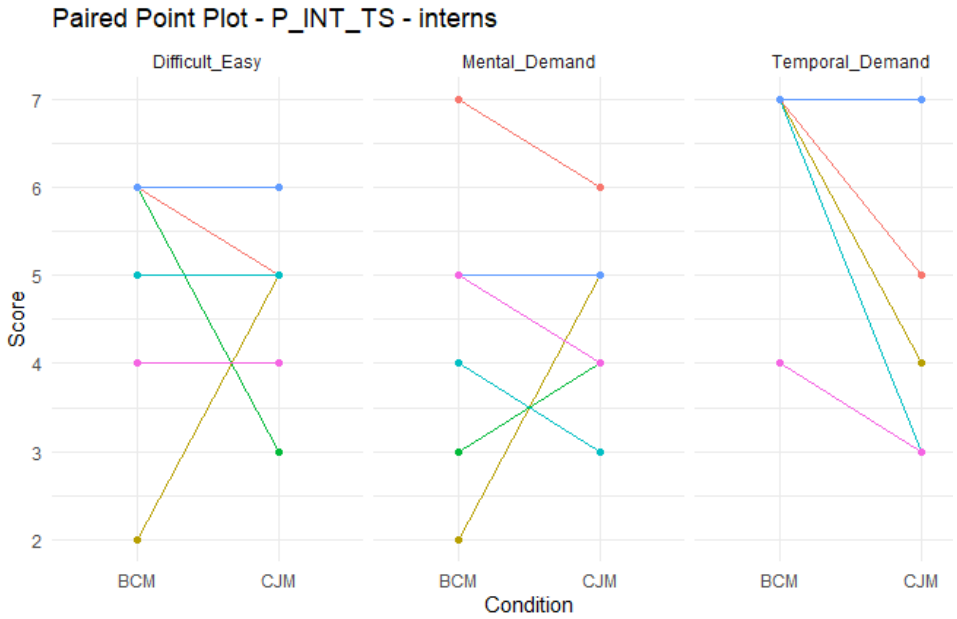
As expected, no significant difference was experienced between the interns and the benchmark because, for this group, the nature of the exercise didn't change.

4.2.3 Task Specific Difficulty

4.2.3.1 Starting Participants (interns)

4.2.3.1.1 Task-Specific feedback

BCM	CJM
No	No
My topic is still very vague / needs to get approved so I am in many doubts	;)
There are some elements I just have not thought about before this task with this perspective.	This one was a bit more difficult because it maybe was not completely applicable to my situation. Also, a bit repetitive at other aspects.
I don't have any.	Not necessarily hurried and rushed but I haven't given too much thought about what I would have to do today and I need to think about it which makes me lose time.
N/A	N/A
NE	Due to time I had to think quicker



4.2.3.1.2 Internal Task Comparison

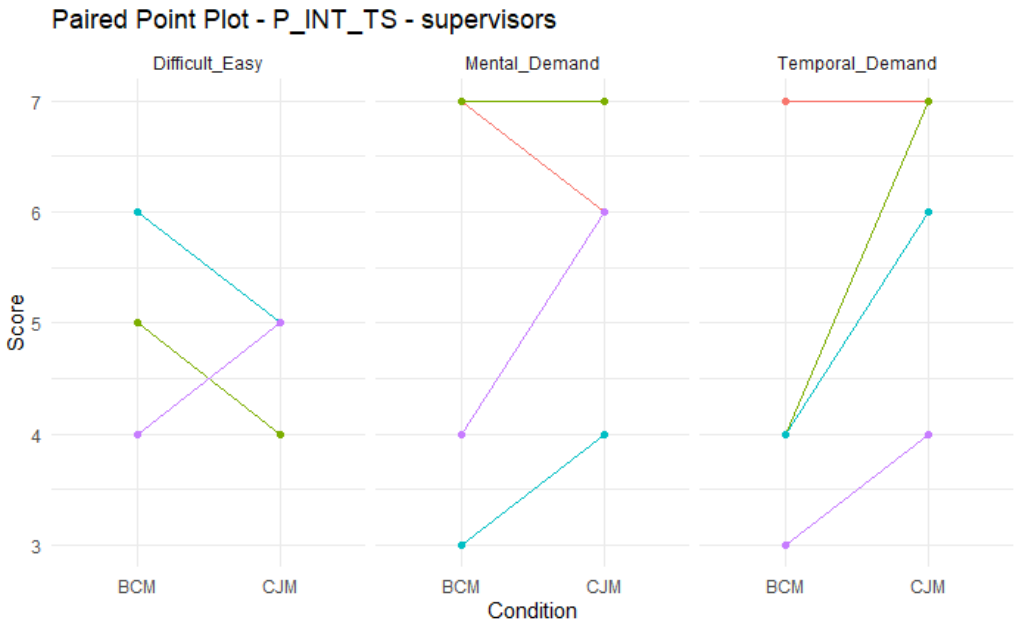
Since the interns and the supervisors were given a different exercise, the analysis also differentiates between them in the internal comparison. The interns consider the BCM and CJM as equally difficult and mentally demanding. The participants consistently experienced the BCM as more rushed and hurried than the CJM. This indicates that the timing of the exercise could be improved to achieve balance.

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	BCM	CJM	4,83	4,67	0,17	1,000	0,190
Mental Demand	BCM	CJM	4,33	4,50	-0,17	1,000	0,048
Temporal Demand	BCM	CJM	6,50	4,17	2,33	0,058	0,952

4.2.3.2 Follow-up Participants (supervisors)

4.2.3.2.1 Task specific feedback

BCM	CJM
-	-
.	Bit more difficult compared to previous assignment because the thesis process is not fully set in stone yet.
As Julia has just started, hew whole idea is still quite new and not well defined. Therefore it was more mentally demanding then for example over a few weeks.	n/a
Limited time because of client work :)	nvt



4.2.3.2.2 Internal Task Comparison

Like the interns, the supervisors experienced no significant difference between the difficulty and mental demands of the exercises. However, surprisingly, they experienced the CJM as more rushed and hurried, contrasting with the interns' experience. This difference does not fully meet the significance threshold, but a strong effect-size difference suggests that this should be considered in the re-design.

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	BCM	CJM	5,25	4,75	0,50	0,424	0,500
Mental Demand	BCM	CJM	5,25	5,75	-0,50	0,586	-0,400
Temporal Demand	BCM	CJM	4,50	6,00	-1,50	0,181	-0,900

In conclusion, **providing more time for the BCM when first filling in the template would be advised. Conversely, it would be good to distribute the time more evenly between the BCM and CJM during the Reviewing interviews.**

4.2.4 User Experience Questionnaire

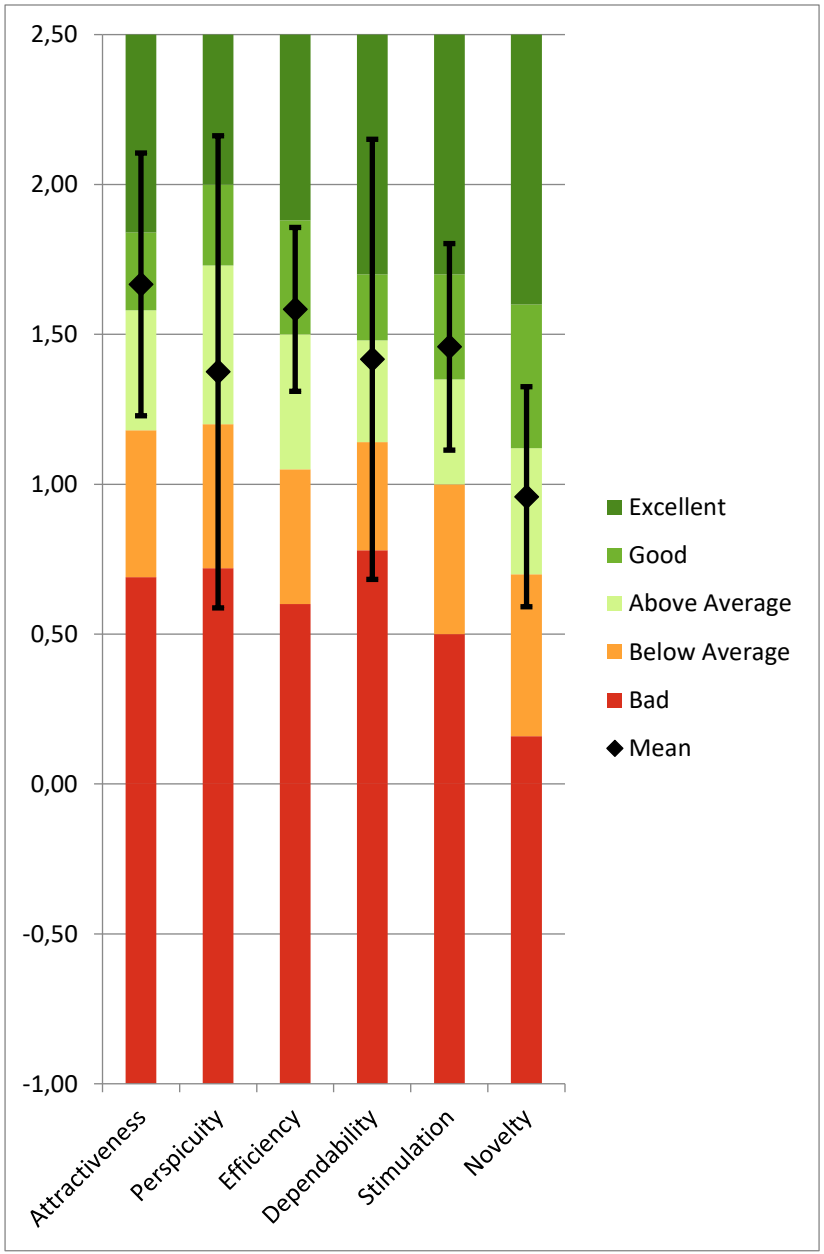
4.2.4.1 P-UEQ – Interview - Interns

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,67	**Good**	10% of results better, 75% of results worse
Perspicuity	1,38	**Above Average**	25% of results better, 50% of results worse
Efficiency	1,58	**Good**	10% of results better, 75% of results worse
Dependability	1,42	**Above Average**	25% of results better, 50% of results worse
Stimulation	1,46	**Good**	10% of results better, 75% of results worse
Novelty	0,96	**Above Average**	25% of results better, 50% of results worse

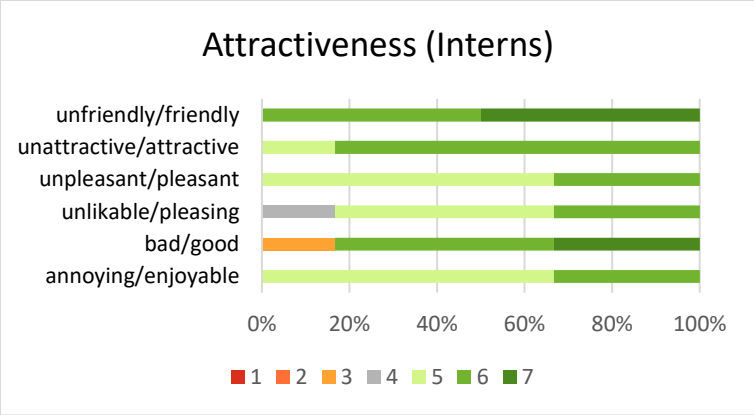
The UEQ scores came back quite positively, showing 'above average' results for all six subscales and even suggesting that half of them performed 'good'. The most notable measure is the perspicuity subscale, which seems to be closest to the benchmark average.

UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,667	0,3	0,438	0,82
Perspicuity	1,375	0,97	0,788	0,69
Efficiency	1,583	0,12	0,273	0,16
Dependability	1,417	0,84	0,734	0,57
Stimulation	1,458	0,19	0,345	0,46
Novelty	0,958	0,21	0,367	0,66

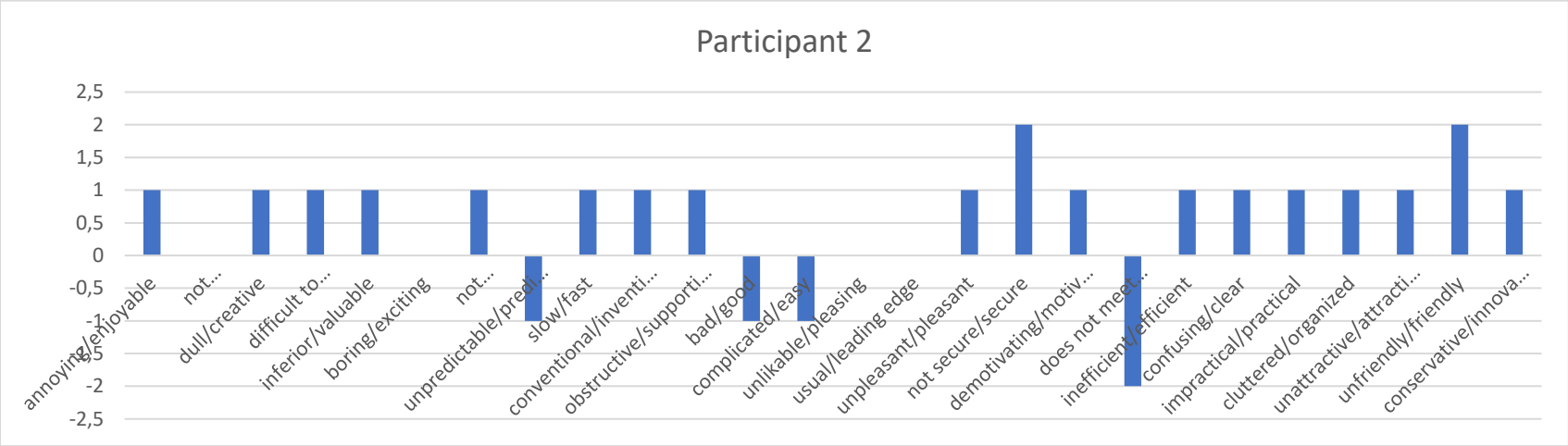
Due to the small sample size, these results should be carefully considered. The efficiency, dependability, and stimulation subscales score poorly on the Alpha scores, which suggests that the opinions were divided on some items in their subscales.

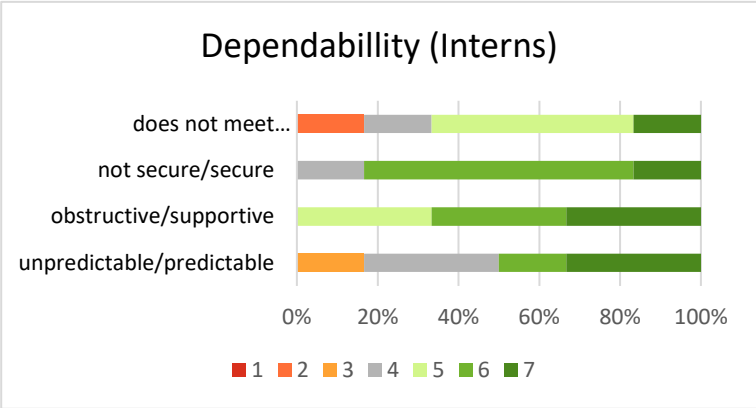


Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,3	0,3	0,5	6	annoying	enjoyable	Attractiveness
2	1,3	1,1	1,0	6	not understandable	understandable	Perspicuity
3	1,7	0,3	0,5	6	creative	dull	Novelty
4	1,3	5,5	2,3	6	easy to learn	difficult to learn	Perspicuity
5	1,8	0,6	0,8	6	valuable	inferior	Stimulation
6	0,8	0,6	0,8	6	boring	exciting	Stimulation
7	1,7	0,3	0,5	6	not interesting	interesting	Stimulation
8	1,2	3,0	1,7	6	unpredictable	predictable	Dependability
9	1,3	0,3	0,5	6	fast	slow	Efficiency
10	0,8	1,0	1,0	6	inventive	conventional	Novelty
11	2,0	0,8	0,9	6	obstructive	supportive	Dependability
12	1,8	2,2	1,5	6	good	bad	Attractiveness
13	1,2	1,4	1,2	6	complicated	easy	Perspicuity
14	1,2	0,6	0,8	6	unlikable	pleasing	Attractiveness
15	-0,2	0,2	0,4	6	usual	leading edge	Novelty
16	1,3	0,3	0,5	6	unpleasant	pleasant	Attractiveness
17	1,8	1,0	1,0	6	secure	not secure	Dependability
18	1,5	0,7	0,8	6	motivating	demotivating	Stimulation
19	0,7	2,7	1,6	6	meets expectations	does not meet expectations	Dependability
20	1,7	0,7	0,8	6	inefficient	efficient	Efficiency
21	1,7	1,1	1,0	6	clear	confusing	Perspicuity
22	1,3	0,3	0,5	6	impractical	practical	Efficiency
23	2,0	0,4	0,6	6	organized	cluttered	Efficiency
24	1,8	0,2	0,4	6	attractive	unattractive	Attractiveness
25	2,5	0,3	0,5	6	friendly	unfriendly	Attractiveness
26	1,5	0,7	0,8	6	conservative	innovative	Novelty



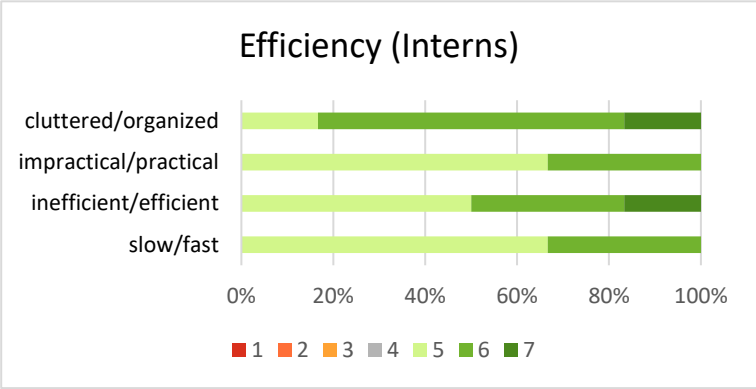
The reason why one of the participants was significantly more negative than the others about the exercise is not clear. It appears that this participant found the exercise complicated and that they had other expectations for this exercise. This participant also expressed that *"My topic is still very vague / needs to get approved so I am in many doubts",* and that *"It would be easier to do a bit further in the process, but I guess some people are much further in their topic".* We can conclude that, in this case, the score can be explained by the early phase of their project, which caused the participant to experience this exercise as more complicated and as less good than others experienced it.



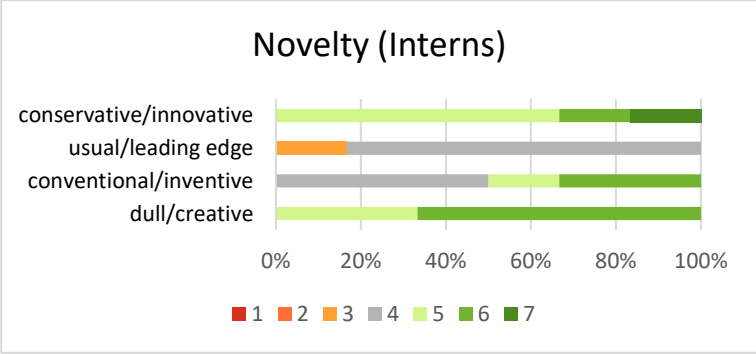


Some did not meet Predictability and expectations; however, it does not state whether this was positive or negative. The 2 and 3 scores are linked to the participant who had yet to define their project, making it likely more difficult to perform this exercise, as they are asked to predict the flow of their project. As such, we assume that this participant was unprepared for the exercise and expected something else.

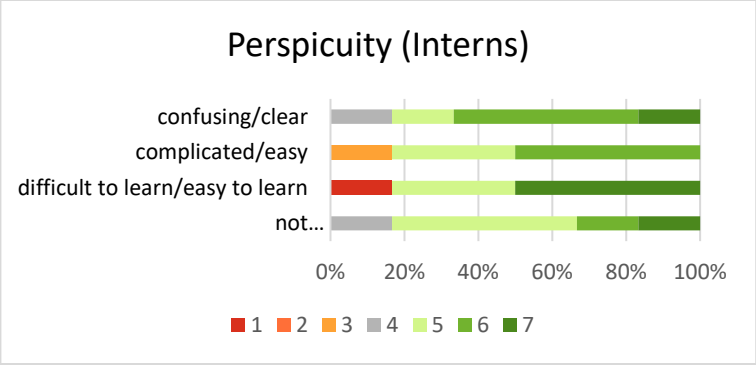
INSIGHT: The context and its development phase can strongly influence the participants' user experience. The context in other sessions will likely be very different and much more evolved. This could make it easier; however, as projects get more complicated, this is expected to negatively impact the user experience.



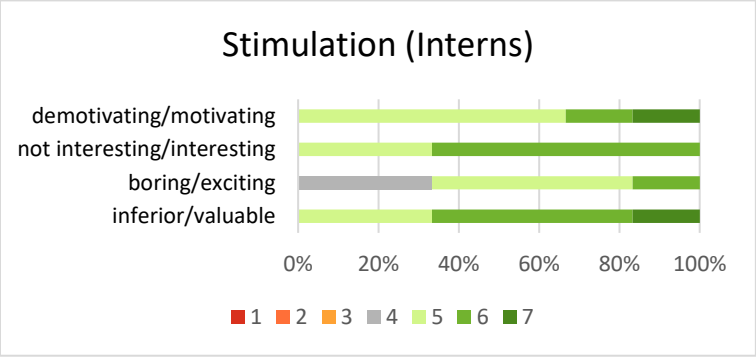
Participants seem to be consistently positive about the efficiency of the exercise.



The exercise is seen as usual and not very cutting-edge, which is not surprising as the interview uses popular templates to capture the data. Regardless, the overall Novelty score is high because the goal and nature of the exercise are focused on innovation and leaning on the participants' creative capabilities.



The perspicuity scores were relatively low, as some participants found the exercise complicated and difficult to learn. Regardless of the opinions of others, it is good to understand why participants experienced this exercise. It might be due to the nature of their project or context, if this is yet undefined, they might experience this exercise as more complicated. As this is difficult to improve, various tips and observations can be utilized to improve the exercise's clarity and overall perspicuity.



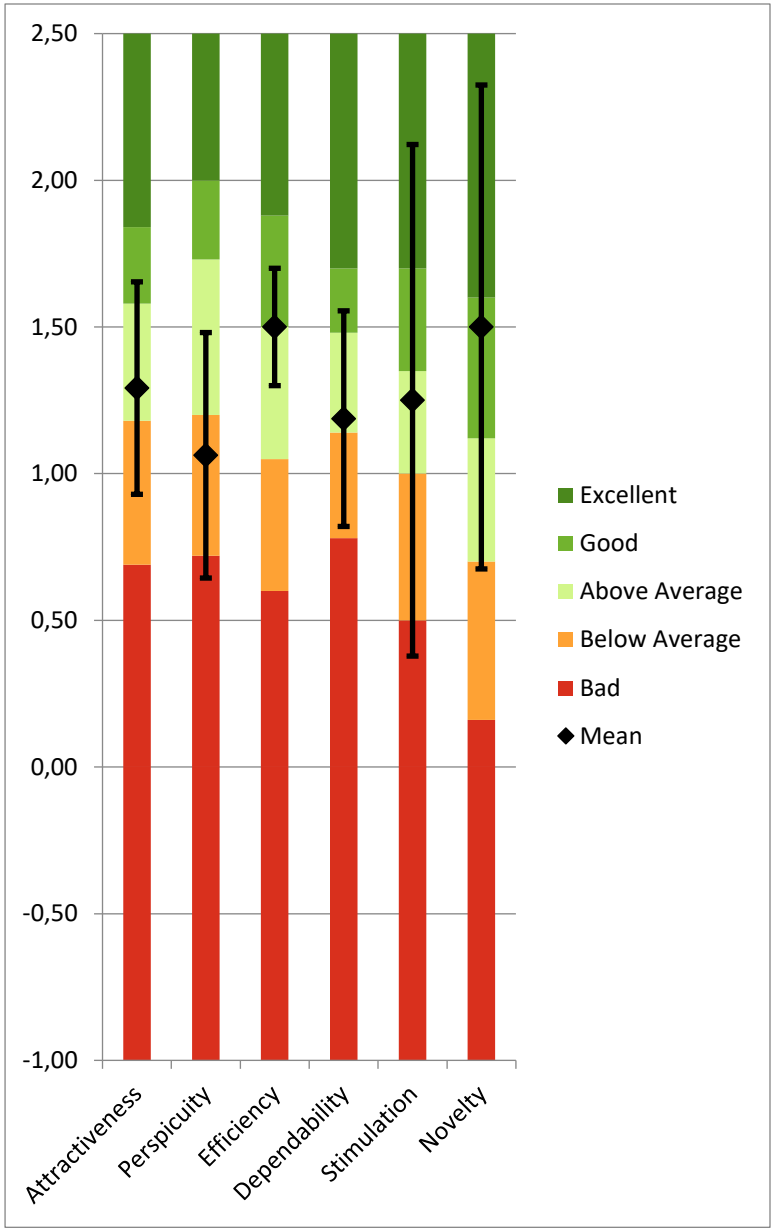
The interns perceived the overall exercise as quite stimulating, although some found it a bit more boring than others. A remark was made on some repetitive elements in the BCM and CJM. By explaining this link, future participants might hopefully find the experience less repetitive and boring.

4.2.4.2 P-UEQ – Interview - Supervisors

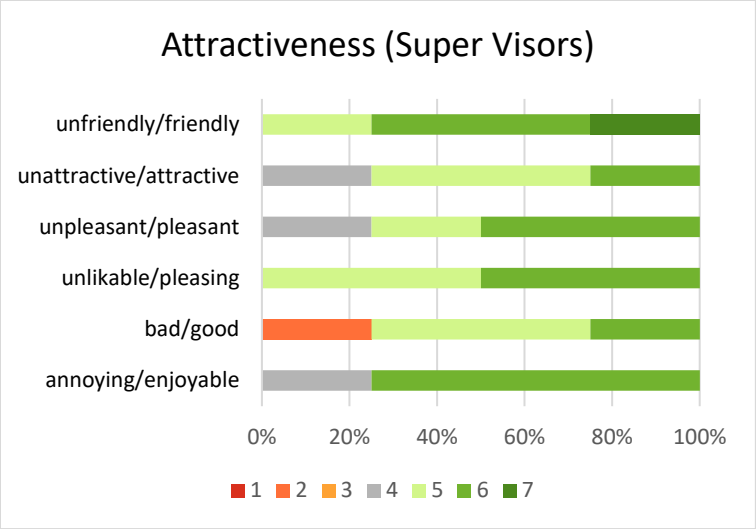
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,292	0,14	0,362	0,4
Perspicuity	1,063	0,18	0,418	-8,29
Efficiency	1,5	0,04	0,200	-1,02
Dependability	1,188	0,14	0,367	-3,69
Stimulation	1,25	0,79	0,872	0,87
Novelty	1,5	0,71	0,825	0,91

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,29	**Above average**	25% of results better, 50% of results worse
Perspicuity	1,06	**Below Average**	50% of results better, 25% of results worse
Efficiency	1,50	**Above Average**	25% of results better, 50% of results worse
Dependability	1,19	**Above Average**	25% of results better, 50% of results worse
Stimulation	1,25	**Above Average**	25% of results better, 50% of results worse
Novelty	1,50	**Good**	10% of results better, 75% of results worse

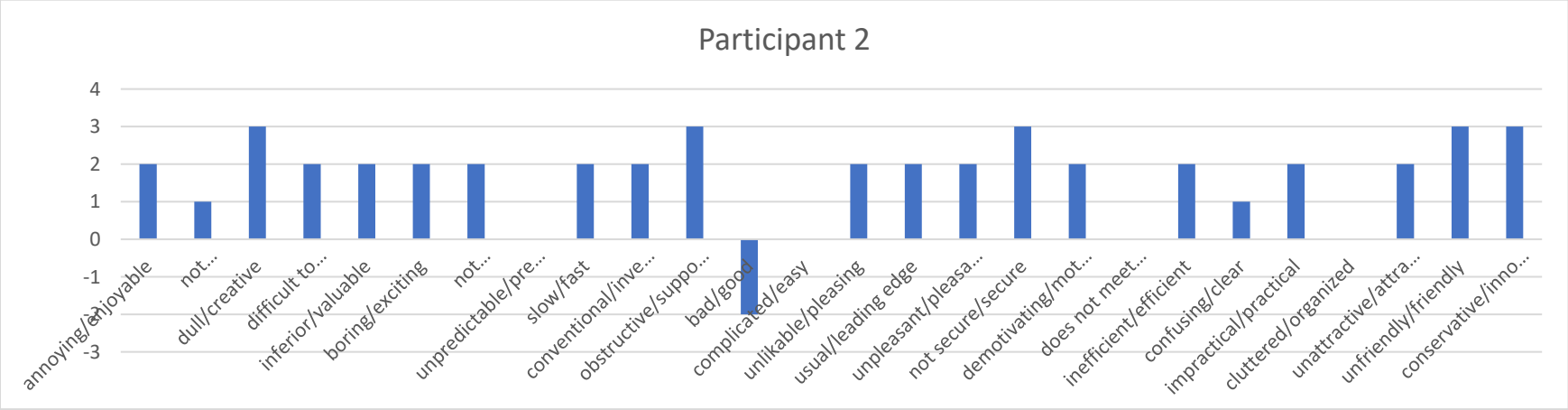
The below-average score here is notable and must be further explored.

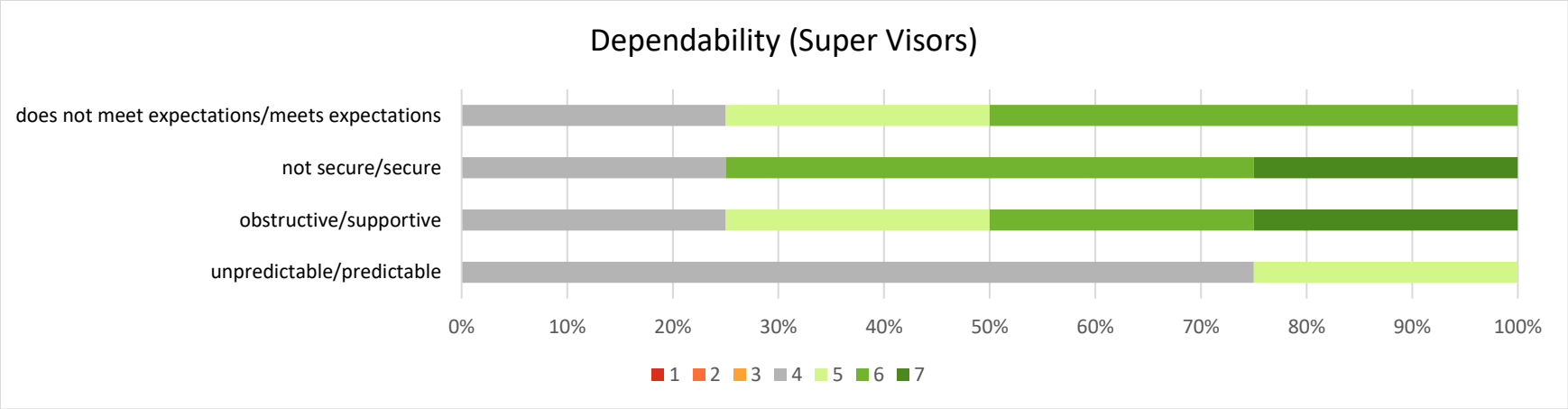


Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,5	1,0	1,0	4	annoying	enjoyable	Attractiveness
2	1,0	0,0	0,0	4	not understandable	understandable	Perspicuity
3	2,3	0,9	1,0	4	creative	dull	Novelty
4	2,0	0,7	0,8	4	easy to learn	difficult to learn	Perspicuity
5	1,8	0,9	1,0	4	valuable	inferior	Stimulation
6	1,0	0,7	0,8	4	boring	exciting	Stimulation
7	1,0	2,0	1,4	4	not interesting	interesting	Stimulation
8	0,3	0,3	0,5	4	unpredictable	predictable	Dependability
9	1,0	0,7	0,8	4	fast	slow	Efficiency
10	1,5	0,3	0,6	4	inventive	conventional	Novelty
11	1,5	1,7	1,3	4	obstructive	supportive	Dependability
12	0,5	3,0	1,7	4	good	bad	Attractiveness
13	0,8	0,9	1,0	4	complicated	easy	Perspicuity
14	1,5	0,3	0,6	4	unlikable	pleasing	Attractiveness
15	0,8	1,6	1,3	4	usual	leading edge	Novelty
16	1,3	0,9	1,0	4	unpleasant	pleasant	Attractiveness
17	1,8	1,6	1,3	4	secure	not secure	Dependability
18	1,3	0,9	1,0	4	motivating	demotivating	Stimulation
19	1,3	0,9	1,0	4	meets expectations	does not meet expectations	Dependability
20	1,8	0,3	0,5	4	inefficient	efficient	Efficiency
21	0,5	3,0	1,7	4	clear	confusing	Perspicuity
22	1,8	0,3	0,5	4	impractical	practical	Efficiency
23	1,5	1,7	1,3	4	organized	cluttered	Efficiency
24	1,0	0,7	0,8	4	attractive	unattractive	Attractiveness
25	2,0	0,7	0,8	4	friendly	unfriendly	Attractiveness
26	1,5	1,0	1,0	4	conservative	innovative	Novelty

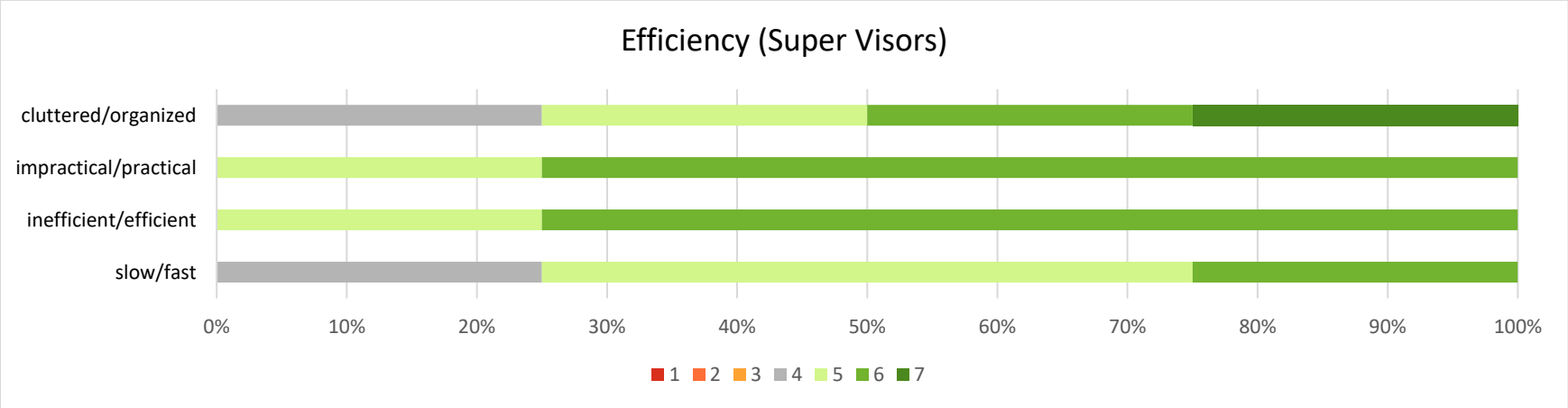


The 2-score on the good/bad scale is quite exceptional and contrasts with the scores from other participants. When looking at these participant-transformed scales, it is surprising that number 12 (Bad/Good) is so strongly contrasting with the rest of the scores. It remains unclear why this score was given and might be considered a fluke.

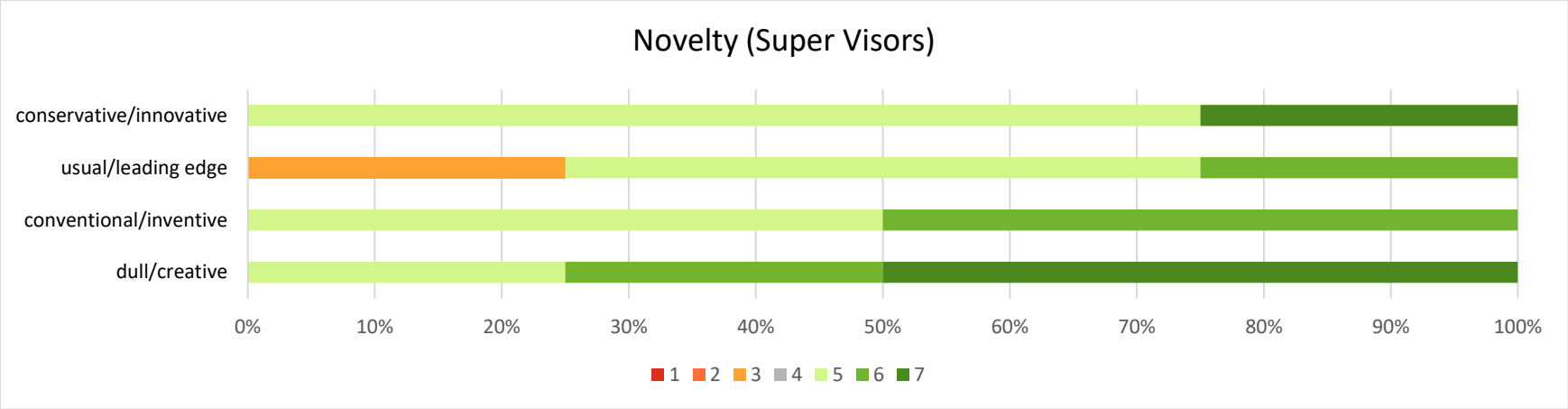


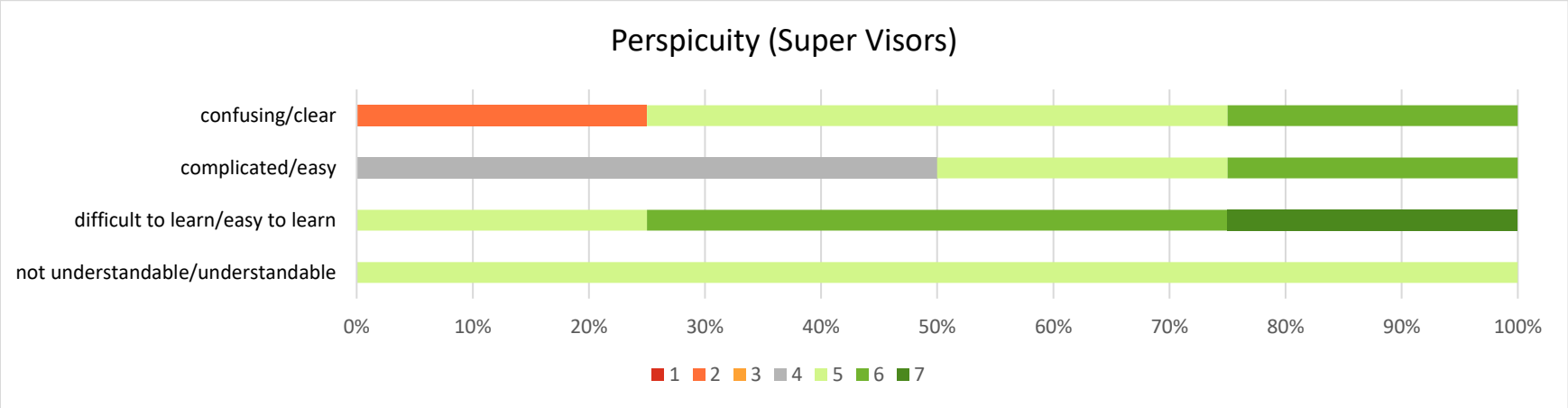


The dependability scores are somewhat lower as the exercise is intended to elicit some unpredictable elements. It will be interesting to see how this score behaves in the next cycle, as people could be surprised by their peers' perspectives but also reinforce their existing mental models.

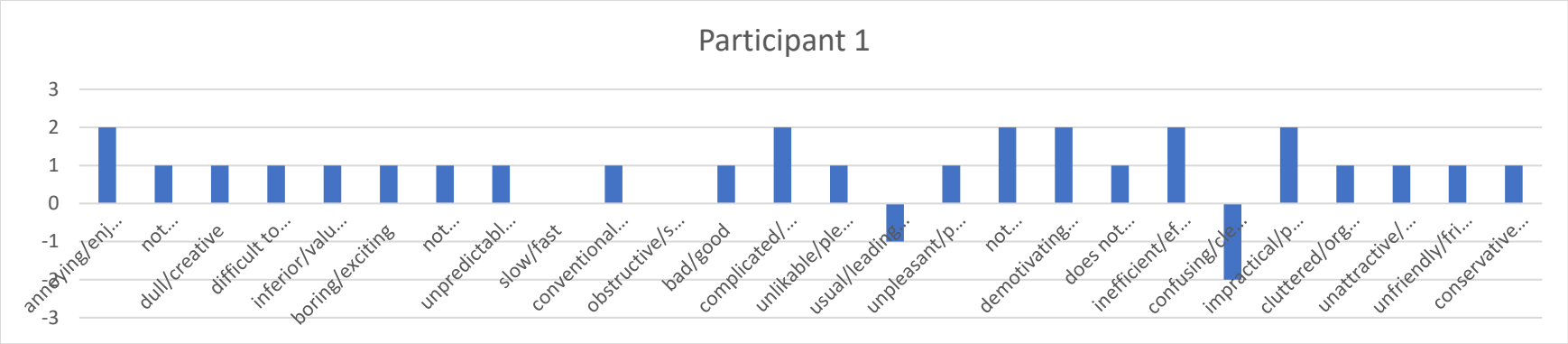


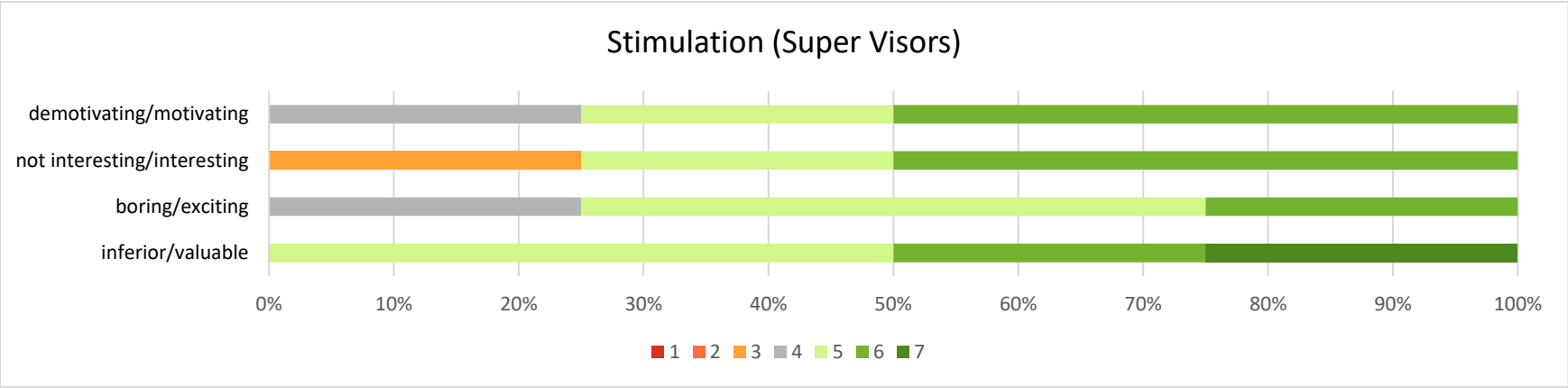
Decluttering the input from others could improve efficiency somewhat. Ensuring readability by using only 2-4 words per Post-it and cleaning the template a little before moving to the follow-up interview could also improve the exercise a bit.





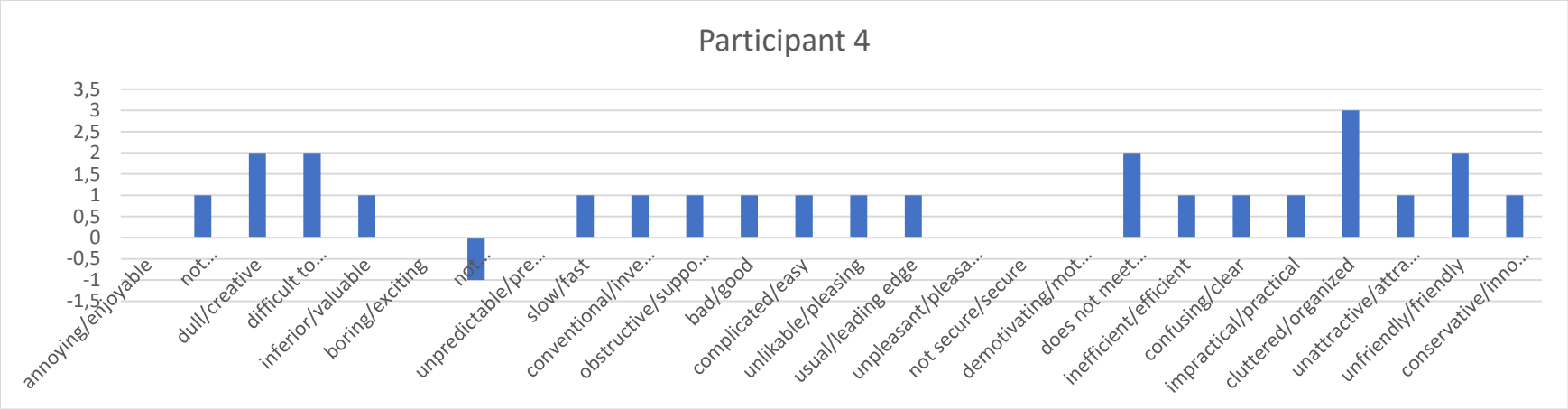
The low scores for Novelty & Perspicuity are both related to one participant. It is not likely that they are related, but they are the only two low scores this participant has given. It must be noted that this was the first supervisor that was interviewed. Based on this experience, some small changes were made to the process that immediately improved the exercise flow (like preparing and pre-tagging the pots-its). Next to this, this participant had a background that is highly comparable to the background of the researcher; this might influence the usual/leading edge scale as this participant is likely more familiar with the templates that are being used. The need for more clarity was explained by the participant as *"put the agenda of the session in the Miro board - make sure the post-its are ready for the people to use in the correct color." On the other hand,* this participant also complemented the exercise by stating, *"Clearly explained and there was enough time.*





The 3 and 4 scores of the stimulation scale all relate back to participant 4. This participant's situation might have influenced this scale, as this participant was onboarding a new, large project: “Limited time because of client work :).” This might have influenced the experience of this participant, as they were less interested in the intern project and more focused on their own work. The participant complimented the exercise: “Very well prepared, good facilitation. All in all, a well-done exercise.”

Overall, this exercise seems to have stimulated participants well. However, it is important to remember that the personal situation and mindset of participants outside of the exercise are likely to influence the results.



Overall, the exercise was well received, and the most negative scores were, to some extent, explainable and understandable. Comparing the Intern and Supervisor scores shows that the interns have an overall higher score than the supervisors. This might be due to their level of involvement; being more intrinsically motivated could influence the overall user experience. It might also be that the interns are more forgiving and that supervisors tend to be more critical.

Overall, the most important focus on improving the design is in **perspicuity**, where both the interns and supervisors score relatively low against the benchmark. By focusing on **improving the clarity and understandability** of the exercise explanations, the hope is that the exercise becomes easier to learn and perform.

4.3 Opportunity Identification Session

4.3.1 End of Survey Comments

Do you have any tips?	Do you have any tops?	Other feedback?
No	No	
No	Great session management	
Miss the connection between the first exercise of circling topics and the steps after	I like to setup of the plastic plates and post its	-
The first step was a bit vague. Feels like whoever is most vocal determines the direction the workshop goes into.	No	That's it. Well done!
Board had quite a lot of input in the beginning, which made getting a clear overview hard	Not enough pencils, thus you don't have all the input. Also, too many people mean you can't reach/read everything	Output was good
Make the end results work	Nice work	
Take a longer and more clear time to introduce the entire workshop before the start. The beginning was rushed.	We actually got some good ideas	
No	No	

Based on the feedback from the participants, the pilot opportunity identification session could benefit from various improvements. Primarily, attention should be paid to the first step in the process. This step embodies the circling of areas of interest on the extensive blueprint. This step intends for participants to familiarize themselves with the blueprint while reflecting on items that might be interesting. However, The feedback is that the circled areas are not included in the following steps, making their purpose unclear. Next to this, the blueprint was perceived as a lot of input, making it harder to get a clear overview in the beginning. To improve the design of the process, the workshop steps and their purpose should be clearly defined before engaging in the exercise. Next to this, it can be considered to exclude the circled areas or have them reoccur in the next step as a starting point.

4.3.2 NASA Task Load Index

4.3.2.1 P – TLX - Opportunity Identification – Pilot vs Bench

TLX Metric	Mean P-OI	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,1	6,8	-2,6	4,5	0,056	0,719
Performance	7,3	7,3	0,0	15,5	1,000	0,031
Mental Demand	3,9	7,0	-3,1	3,0	0,030	0,813
Temporal Demand	3,6	6,3	-2,6	6,0	0,097	0,625
Frustration	1,9	1,0	0,9	19,0	0,647	-0,188

4.3.3 Task Specific Difficulty

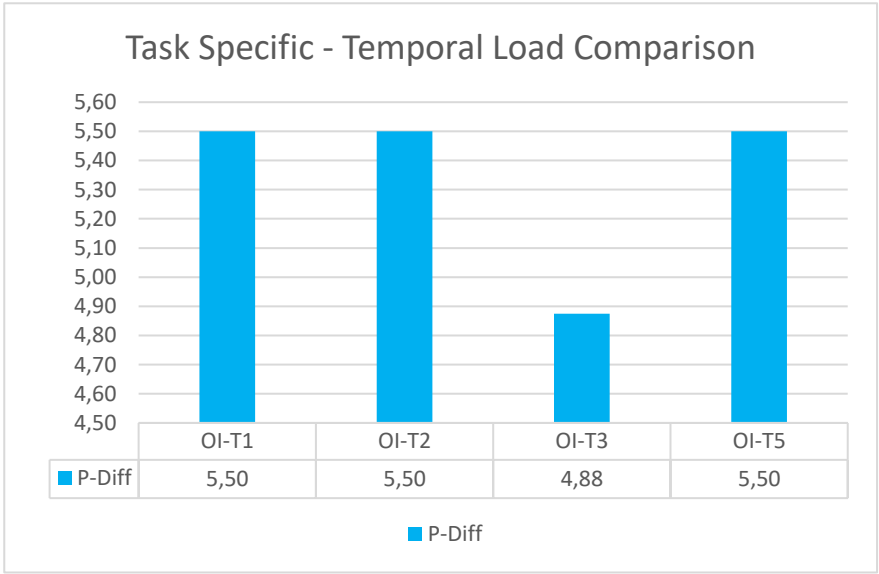
4.3.3.1 Task specific feedback

T1	T2	T3	T4	T5
No measure below 4	No measure below 4	No measure below 4	No measure below 4	No measure below 4
N/a	N/a	N/A	N/A	N/A
-	-	I have done this exercise already this week, so I feel like I don't have any ideas anymore	-	-
N/a	N/a	N/a	N/a	N/a
A lot of items are already on the board so easy to map	N/a	Ok	The answers were more based on what I could physically reach on the board	Ok
No	No	No	No	No
It takes very long to read the entire map	No	Im not creative	.	.
The conversation was dominated by supervisors	Same as previous comment	N/A	N/A	N/A

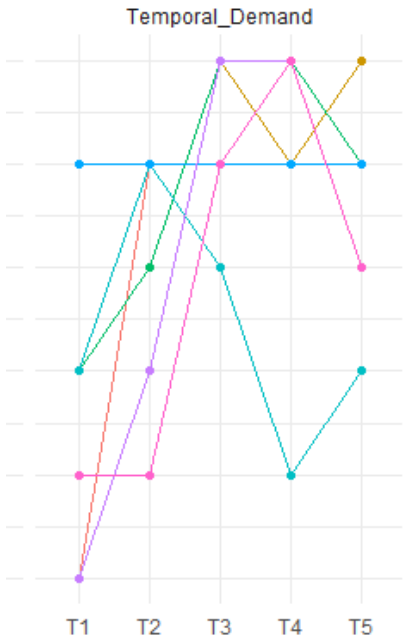
4.3.3.2 Internal Task Comparison

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	T1	T2	5,50	5,50	0,00	0,95	0,03
Difficulty	T1	T3	5,50	4,88	0,63	0,38	-0,27
Difficulty	T1	T4	5,50	5,63	-0,13	0,78	0,09
Difficulty	T1	T5	5,50	5,50	0,00	1,00	0,00
Difficulty	T2	T3	5,50	4,88	0,63	0,35	-0,27
Difficulty	T2	T4	5,50	5,63	-0,13	0,82	0,08
Difficulty	T2	T5	5,50	5,50	0,00	0,86	-0,06
Difficulty	T3	T4	4,88	5,63	-0,75	0,26	0,33
Difficulty	T3	T5	4,88	5,50	-0,63	0,39	0,25
Difficulty	T4	T5	5,63	5,50	0,13	0,69	-0,13

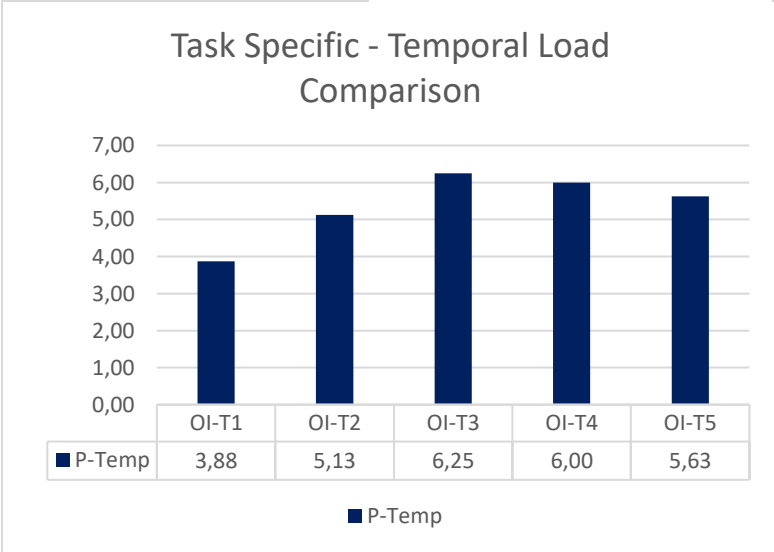
No significant differences in difficulty were found when comparing the tasks internally. The biggest difference is found between task 3 and tasks 4 and 5, in which task 3 was evaluated as slightly more difficult than the other two tasks. This is likely because task three requires the participants to identify new opportunities, which requires a creative component that is not part of tasks 4 and 5 and was mentioned as a reason in the comments.



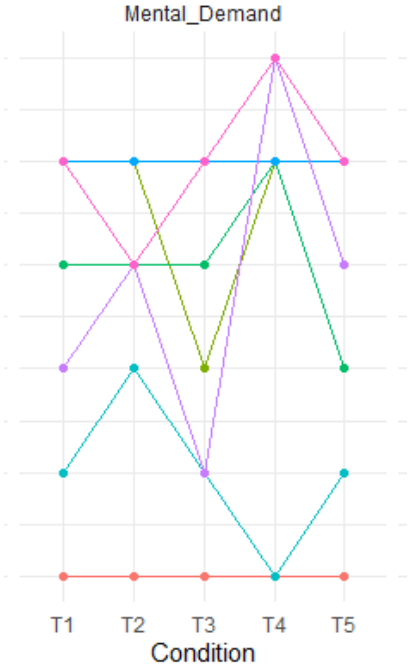
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Temporal Demand	T1	T2	3,88	5,13	-1,25	0,11	0,47
Temporal Demand	T1	T3	3,88	6,25	-2,38	0,01	0,81
Temporal Demand	T1	T4	3,88	6,00	-2,13	0,02	0,70
Temporal Demand	T1	T5	3,88	5,63	-1,75	0,03	0,64
Temporal Demand	T2	T3	5,13	6,25	-1,13	0,04	0,59
Temporal Demand	T2	T4	5,13	6,00	-0,88	0,07	0,52
Temporal Demand	T2	T5	5,13	5,63	-0,50	0,43	0,23
Temporal Demand	T3	T4	6,25	6,00	0,25	1,00	-0,02
Temporal Demand	T3	T5	6,25	5,63	0,63	0,17	-0,39
Temporal Demand	T4	T5	6,00	5,63	0,38	0,26	-0,33



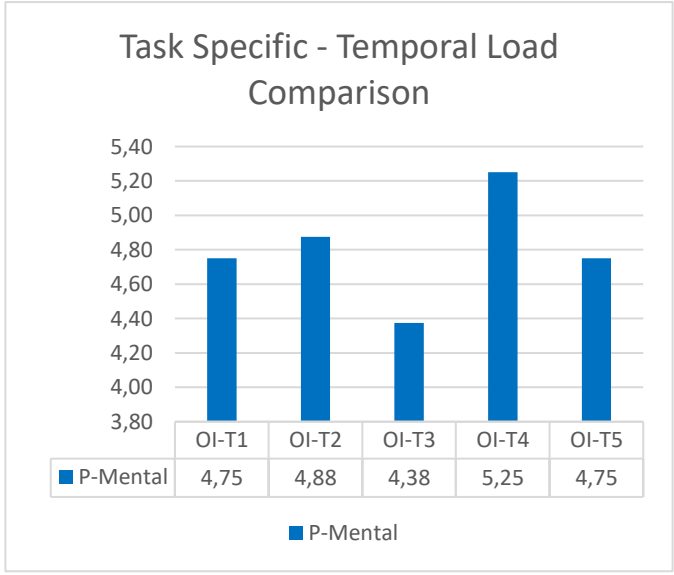
The difference in the difficulty of task three is also seen in the mental demand metric, in which, again, no significant difference is found between the tasks. However, the most significant differences (P < 0,3) again show that task three is more mentally demanding than tasks 2, 4, and 5.



Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Mental Demand	T1	T2	4,75	4,88	-0,13	1,00	-0,02
Mental Demand	T1	T3	4,75	4,38	0,38	0,66	-0,14
Mental Demand	T1	T4	4,75	5,25	-0,50	0,34	0,28
Mental Demand	T1	T5	4,75	4,75	0,00	1,00	0,00
Mental Demand	T2	T3	4,88	4,38	0,50	0,62	-0,16
Mental Demand	T2	T4	4,88	5,25	-0,38	0,25	0,34
Mental Demand	T2	T5	4,88	4,75	0,13	1,00	0,02
Mental Demand	T3	T4	4,38	5,25	-0,88	0,25	0,34
Mental Demand	T3	T5	4,38	4,75	-0,38	0,66	0,14
Mental Demand	T4	T5	5,25	4,75	0,50	0,34	-0,28



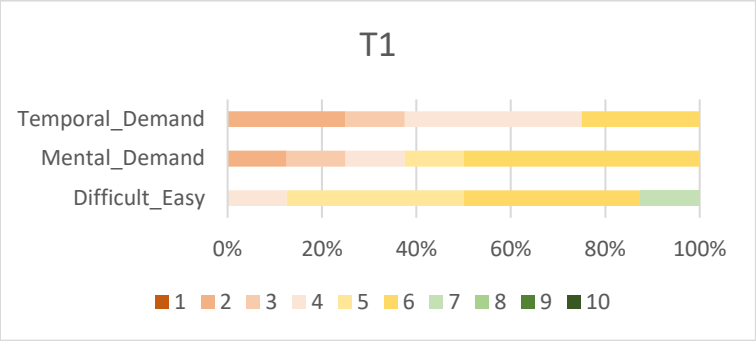
For the temporal demand, we find significant differences when internally comparing tasks. This is mainly due to task 1, which was evaluated as having a high temporal demand in comparison to the other tasks. The comments indicate this difference: the need for more time to digest the information on the blueprint. We can conclude from this that more time should be allocated for the first task; seemingly, this time could be drawn from task 3. However, since we found indications of task 3 being more difficult and mentally demanding than others, the time for this task is likely better relocated from tasks 2, 4, and 5.



Looking at the distribution of answers and the participant evaluation, it first stands out that one participant evaluated the mental demand as very high for all steps in this exercise (scoring it with a value of two for each exercise). Regardless of the survey asking for an explanation, this participant did not explain why he had graded this for any of its measures below 4. Another surprising fact is that one participant gave all the tasks a score of six, which is unlikely to represent their experience.

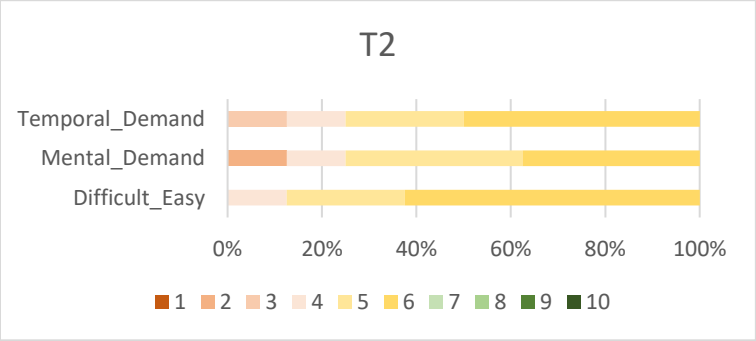
4.3.3.3 Task Specific Data Distribution

4.3.3.3.1 P-TS – Opportunity Identification - Task 1

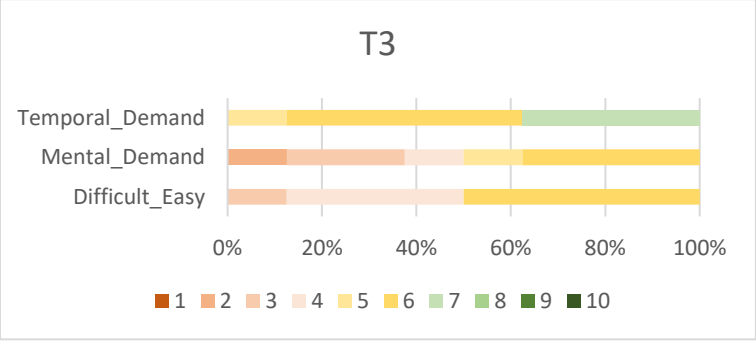


The participants considered the first task of identifying areas of interest by circling them on the blueprint fairly easy but quite mentally demanding, requiring more time to digest the information on the template.

4.3.3.3.2 P-TS – Opportunity Identification - Task 2



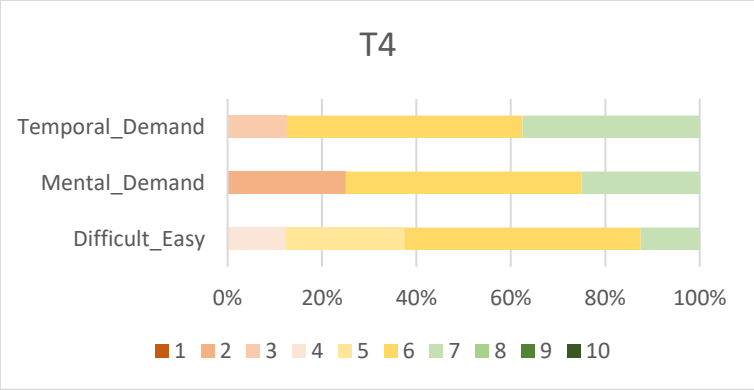
The second task of positioning the existing offerings on the template was again mentally demanding. This is likely still due to participants' need to familiarize themselves with the blueprint. The blueprint should either contain less information or participants should be given more time to familiarize themselves with such a mapping. Another option to reduce the task load of this task is to reduce the number of existing opportunities that the group covers.



4.3.3.3.4 P-TS – Opportunity Identification - Task 3

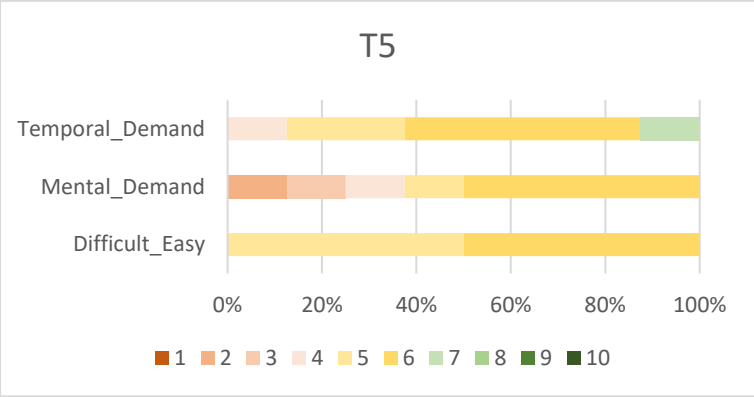
A clear movement in the metrics is seen with task 3 in the paired plot point diagram, as this task is considered by most participants the most difficult and mentally demanding task. One comment suggests that the lack of creativity is due to the difficulty; another suggests that the exercise is comparable to the task of identifying existing synergies efforts during the interview. These were considered very different; however, they might be quite closely related. It might be helpful to explain more clearly during the interview that it should only be about existing synergies and that hypothesizing about potential opportunities is necessary during the follow-up workshop.

4.3.3.3.4 P-TS – Opportunity Identification - Task 4



The pre-selection task was, by many, considered the easiest and least mentally/ temporally demanding task of the workshop. Most likely, they were familiar with the template and only had to focus on the 10-20 opportunities that were noted down in keywords. In practice, however, this exercise required some time as the group had to wait for the last person to finish, which took some time as walking around the blueprint was tricky. Based on this insight, the exercise could be improved by stimulating the participants to walk around the blueprint more to expose them to the full blueprint, as it was, for many, a challenge to physically reach all the blueprint opportunities.

4.3.3.3.5 P-TS – Opportunity Identification - Task 5



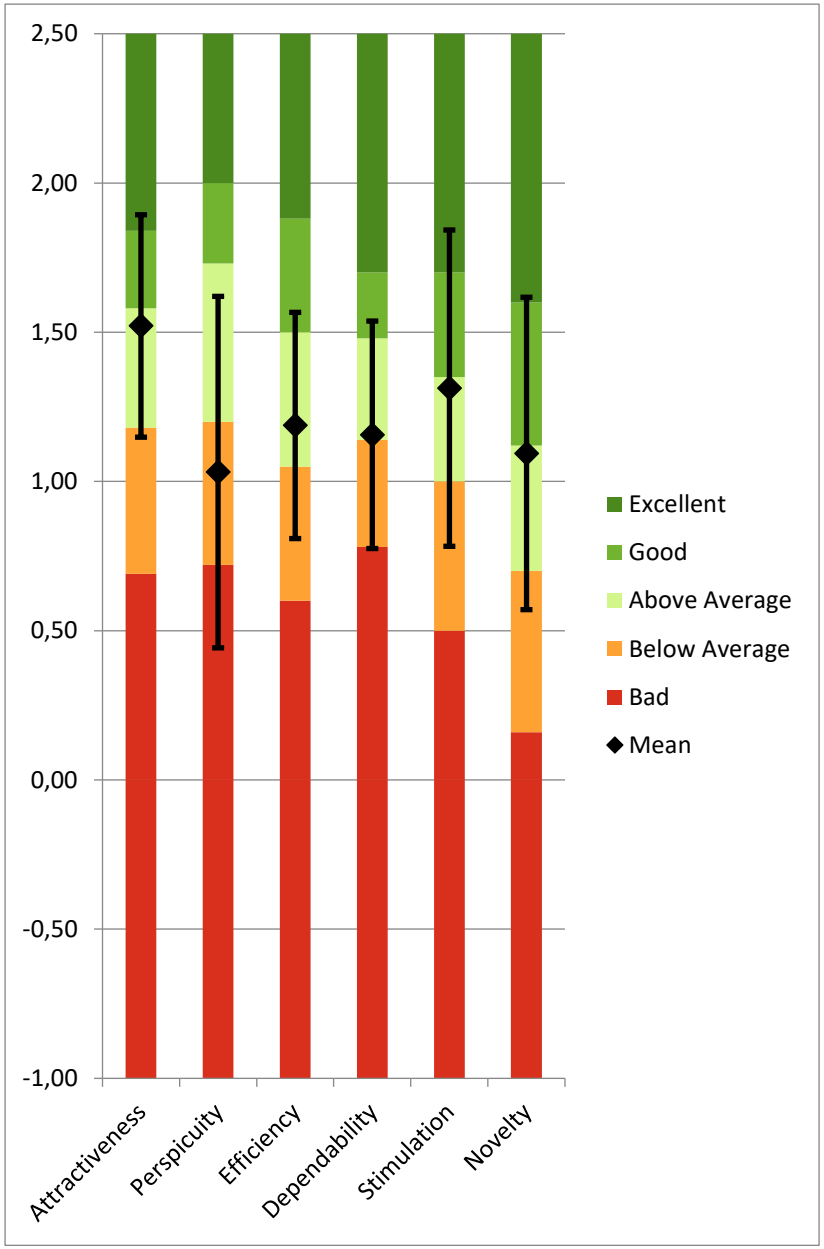
The exercise of framing the opportunity was interesting but was more time-consuming than the designer had intended, as only one opportunity was framed by a group of participants in the allocated time. Based on observation, participants like to create groups with more than two people. This, however, reduces the number of framed opportunities, and many participants took a more passive role, which is not surprising as they are likely to become more tired at the end of the session. To improve the effectiveness of the task, in future workshops, it is recommended that the participants work in pairs.

4.3.4 User Experience Questionnaire

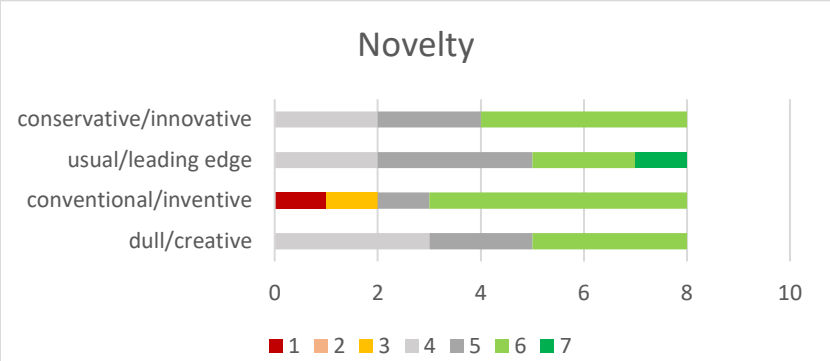
The workshop scored above average overall on the six user experience metrics, except for perspicuity.

UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,521	0,29	0,373	0,83
Perspicuity	1,031	0,72	0,589	0,63
Efficiency	1,188	0,3	0,379	0,11
Dependability	1,156	0,3	0,381	0,21
Stimulation	1,313	0,58	0,530	0,9
Novelty	1,094	0,57	0,523	0,68

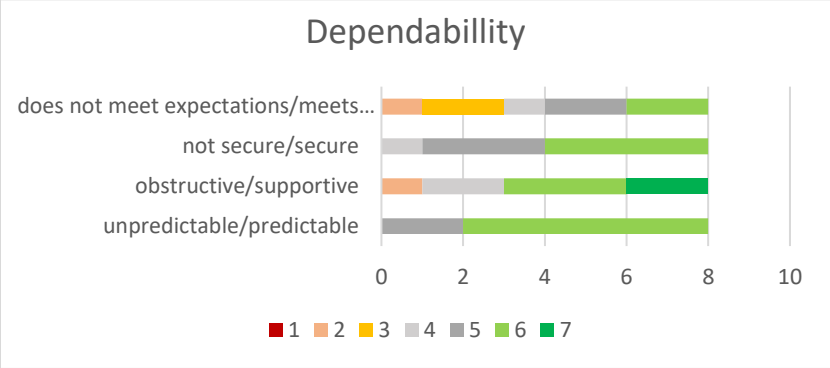
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,52	**Above average**	25% of results better, 50% of results worse
Perspicuity	1,03	**Below Average**	50% of results better, 25% of results worse
Efficiency	1,19	**Above Average**	25% of results better, 50% of results worse
Dependability	1,16	**Above Average**	25% of results better, 50% of results worse
Stimulation	1,31	**Above Average**	25% of results better, 50% of results worse
Novelty	1,09	**Above Average**	25% of results better, 50% of results worse



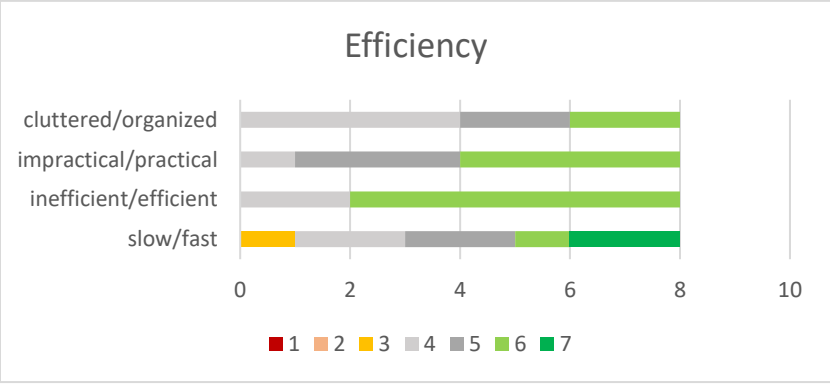
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,8	0,5	0,7	8	annoying	enjoyable	Attractiveness
2	1,1	0,4	0,6	8	not understandable	understandable	Perspicuity
3	0,9	3,6	1,9	8	creative	dull	Novelty
4	1,6	1,4	1,2	8	easy to learn	difficult to learn	Perspicuity
5	1,6	0,8	0,9	8	valuable	inferior	Stimulation
6	0,8	0,8	0,9	8	boring	exciting	Stimulation
7	1,5	0,3	0,5	8	not interesting	interesting	Stimulation
8	0,3	2,2	1,5	8	unpredictable	predictable	Dependability
9	0,8	0,8	0,9	8	fast	slow	Efficiency
10	1,3	1,1	1,0	8	inventive	conventional	Novelty
11	1,8	0,2	0,5	8	obstructive	supportive	Dependability
12	1,9	0,4	0,6	8	good	bad	Attractiveness
13	0,5	1,4	1,2	8	complicated	easy	Perspicuity
14	1,3	0,5	0,7	8	unlikable	pleasing	Attractiveness
15	1,0	0,9	0,9	8	usual	leading edge	Novelty
16	1,1	0,7	0,8	8	unpleasant	pleasant	Attractiveness
17	1,3	3,1	1,8	8	secure	not secure	Dependability
18	1,4	1,4	1,2	8	motivating	demotivating	Stimulation
19	1,4	0,6	0,7	8	meets expectations	does not meet expectations	Dependability
20	1,4	0,6	0,7	8	inefficient	efficient	Efficiency
21	0,9	2,1	1,5	8	clear	confusing	Perspicuity
22	1,5	0,9	0,9	8	impractical	practical	Efficiency
23	1,1	2,1	1,5	8	organized	cluttered	Efficiency
24	1,1	0,7	0,8	8	attractive	unattractive	Attractiveness
25	2,0	0,3	0,5	8	friendly	unfriendly	Attractiveness
26	1,3	0,8	0,9	8	conservative	innovative	Novelty



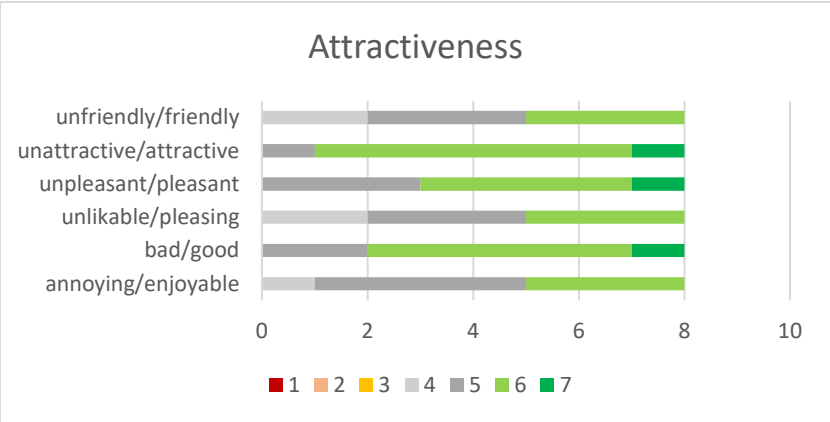
Various comments about the blueprint's design could be considered, such as using a color gradient to reduce the crowded feeling. However, overall, the attractiveness was considered quite good.



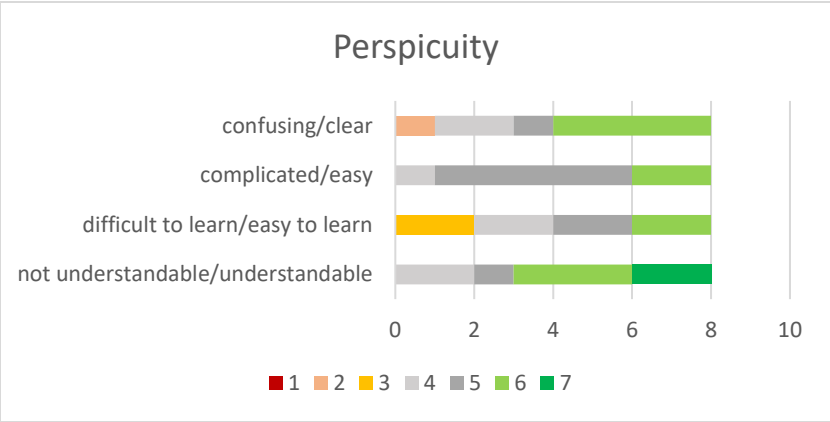
From the dependability metric, it stands out that three participants' expectations are not fully met. There can be various reasons for this. However, it could be improved by being clearer about the nature of the workshop beforehand to manage expectations better.



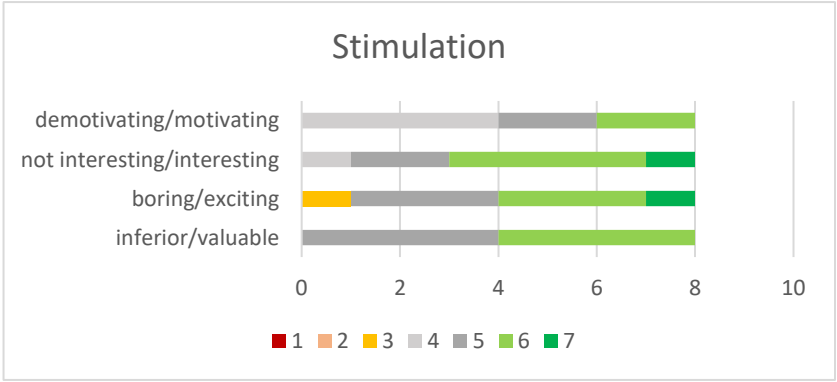
Overall, the workshop was considered quite efficient, organized, and practical. Many found the workshop quite fast; however, one participant considered it to be slow, which is quite contrasting to the rest. The reason behind this, however, is unclear. It might be due to one participant being inpatient or having other priorities than this workshop, but this is purely speculative.



Participants rated the workshop as innovative, leading-edge, and creative. Most were also very inventive; however, two participants found the workshop to be highly conventional. The contrast might indicate an error in the participant's data recording, as left and right differentiated across the survey. Since the participant who voted for highly conventional was an outspoken and critical supervisor, it might also be that this might have been highly conventional in their eyes.



The Perspicuity metric is the only metric that scored below average on the benchmark. To improve this metric, we can, based on these results, target clarity and ease of learning. It is, for instance, likely beneficial to improve the clarity of the introductory explanation of the logical sequence of the tasks, making it easier to learn. Additionally, it might be helpful to explain more clearly what the goal and reason is behind the workshop to ensure everyone is aligned and tries to achieve the same thing. Lastly, the blueprint design could be improved as it was not always clear what specific themes represented; the use of clear headers and text prioritization could be helpful in improving the perspicuity of the blueprint design.



All in all, it was considered quite stimulating, with only one participant finding it moderately boring.

4.4 Opportunity Selection

4.4.1 End of Survey Comments

Do you have any tips?	Do you have any tops?	Other feedback?
No	No	Good session. Possibly pressed on time
No	It was highly interactive	
Not really	No	
One person took over could skew data but leadership is also needed	No	
More time for the sessions	Post-its make it effective	-
Manage the session in terms of who's speaking	Visually good!	It felt a bit repetitive or unnecessary at times
Prioritizing the opportunities after putting them in sequences is less efficient then doing that after the first matrix because you already did the priority there also	Easy to grasp, the quick dynamics make it more challenging (positive way)	
With the survey maybe keep the positive always on the same side	Good clean workshop	
er was meer tijd nodig hiervoor	Structuur was erg duidelijk	

4.4.2 Nasa Task Load Index

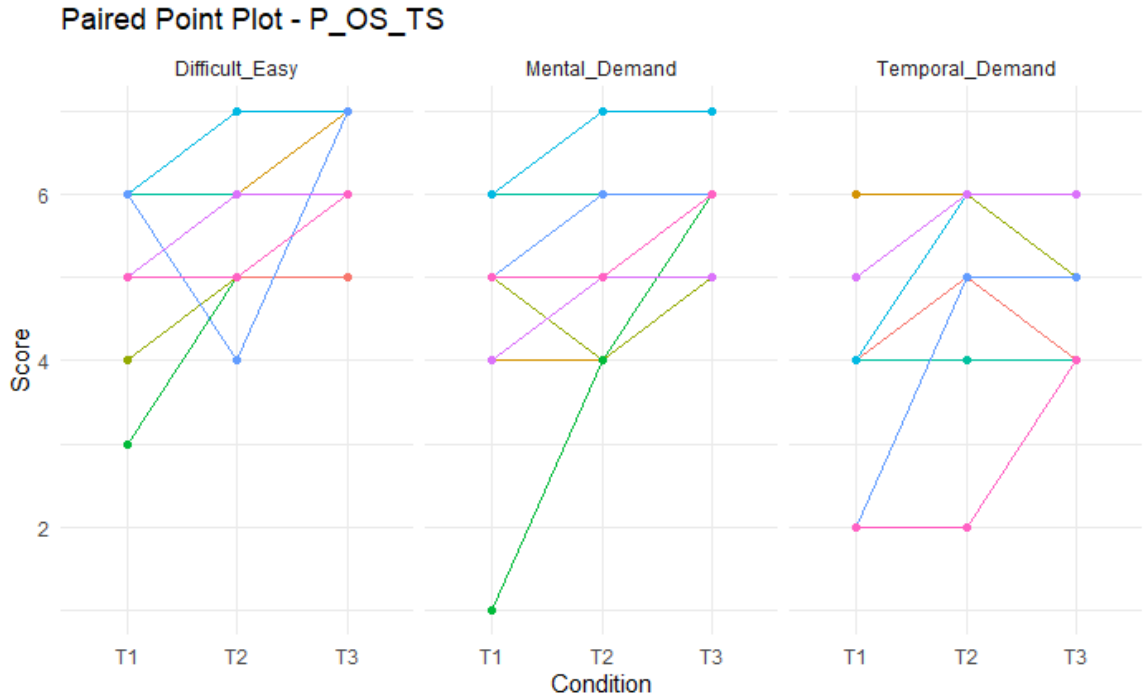
4.4.2.1 P-TLX – Opportunity Selection – Pilot vs Bench

TLX Metric	Mean P-OS	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	5,2	7,5	-2,3	5,5	0,059	0,694
Performance	6,2	7,0	-0,8	17,0	0,936	0,056
Mental Demand	2,3	8,0	1,1	6,5	0,084	0,639
Temporal Demand	5,7	6,3	-0,6	16,0	0,813	0,111
Frustration	6,2	1,3	-0,8	25,0	0,288	-0,389

4.4.3 Task Specific Difficulty

4.4.3.1 Task-Specific Feedback

T1	**T2**	**T3**
N/a	N/a	N/a
N/A	N/A	N/A
No answers below 4	No answers below 4	No lower than 4
I did not really think it was mentally demanding since it was a conjoint activity and the constructs were all understandable although overlapping	No measure below 4	It Was fine
-	-	-
-	No sequence was needed	Feel like we already did this with the first step
Group discussion are more time intensive	Unsure if done right	X
Nog below 4	Niks	.
Te weinig tijd, maar het is wel een goede opdracht. En de duplicates er beter van tevoren uit halen	Komt ook door mij, want ik probeer het af te krijgen binnen de tijd, excuses	Mensen waren het nu sneller eens



4.4.3.2 Internal Task Comparison

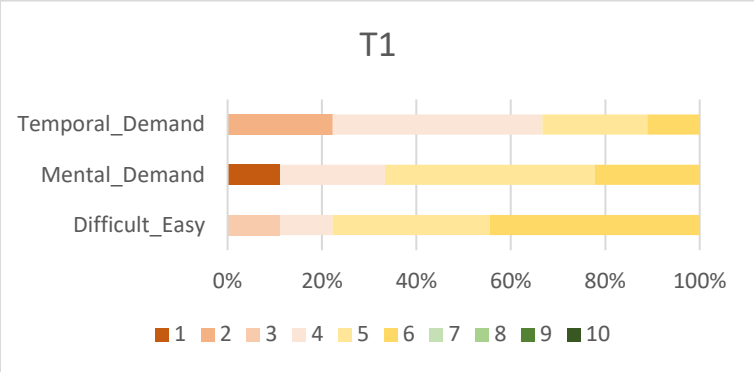
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	T1	T2	5,11	0,54	4,57	0,67	0,12
Difficulty	T1	T3	5,11	0,62	4,49	0,02	0,62
Difficulty	T1	T4	5,44	0,62	4,49	0,06	0,51

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Temporal Demand	T1	T2	4,00	5,11	-1,11	0,06	0,52
Temporal Demand	T1	T3	4,00	5,11	-1,11	0,08	0,48
Temporal Demand	T1	T4	5,11	5,11	0,00	0,74	1,00

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Mental Demand	T1	T2	4,56	5,22	-0,67	0,46	0,21
Mental Demand	T1	T3	4,56	5,89	-1,33	0,01	0,65
Mental Demand	T1	T4	5,22	5,89	-0,67	0,17	0,37

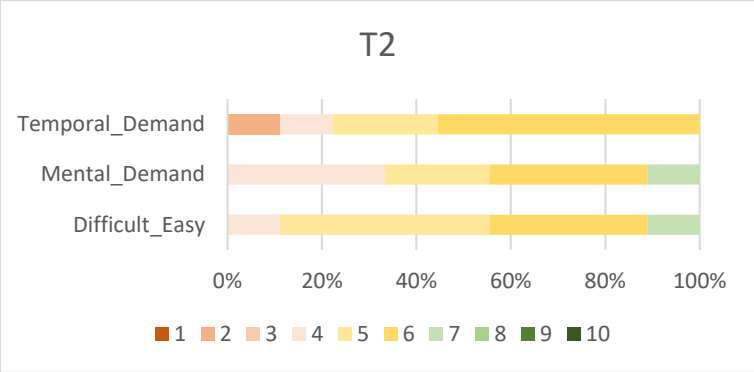
4.4.3.3 Task Specific Data Distribution

4.4.3.3.1 P-TS – Opportunity Selection - Task 1



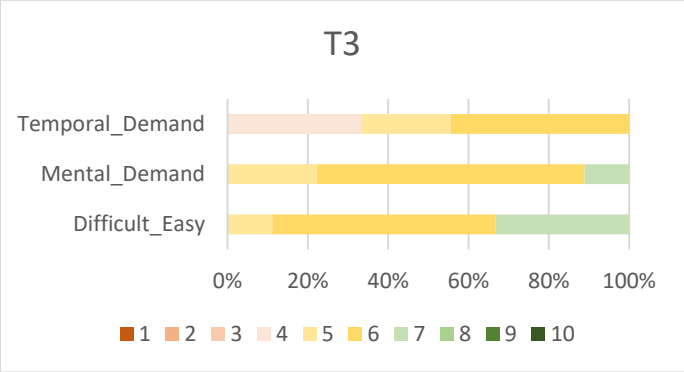
For the first exercise, the group was asked to spend 7 minutes mapping the opportunities based on priority and long/short-term impact. This group was presented with 16 opportunities that were identified during the prior workshop. Participants expressed that these were not properly filtered, causing some duplicates to slow the process down. Next this 7-minute group discussion about positioning 16 opportunities on the matrix could benefit from more time, as many participants experienced a relatively high temporal demand on this exercise. It is expected that if the time per opportunity is increased, the task load for task 1 will improve. Additionally, it is likely helpful if the stack of opportunities is not given to the group but is read out loud and guided by the facilitator to ensure that each perspective is well represented, preventing strongly expressive participants from overshadowing the exercise. An observation is as well that the opportunities were sometimes difficult to interpret, some time had passed, and therefore the participants needed to frame the description to understand the opportunity. It is the question if this mental demand can be further reduced, for instance, by only covering properly framed topics or by linking the opportunity to a general location on the blueprint to support contextual understanding. This last part was, however, attempted by showing a blueprint with positioned opportunities on a screen during the session. However, due to the high pace of the exercise, this was completely ignored.

4.4.3.3.2 P-TS – Opportunity Selection - Task 2



For the second step, the group was provided the choice of getting a new stack of the same opportunities or using the opportunities on the matrix. They chose to use the opportunities on the matrix as they were already familiarized with them and their position on the matrix. The exercise asked them to find sequences among the exercises, which was an exercise that was inspired by literature. The structure on the matrix was dismantled and rearranged into three groups of sequential structures and various separate parts. The task performed better on temporal and mental demand than the first exercise but still was experienced as pressured. Next to this, a participant reported being unsure if the exercise was performed properly, which could indicate unclarity in the explanation. Another participant suggested that this exercise might be unnecessary. The impact on the process, however, was that the group selected intern onboarding over other opportunities, for which the primary reason was that this would influence almost all other opportunities that were identified. Which was identified during the sequencing task.

4.4.3.3.3 P-TS – Opportunity Selection - Task 3

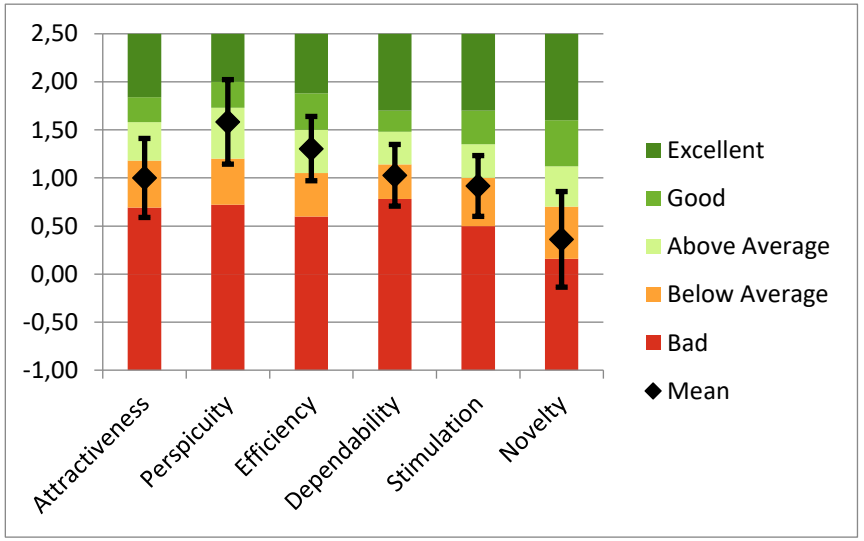


The last exercise was considered easiest among the participants; making a priority list from the opportunities was considered the least mentally demanding and equally rushed as task 2. Participants reported that this exercise felt repetitive as it requested the creation of a priority list, which was also part of the first task. The group only worked out a top 5 list based on time pressure and the intuitive insight that a more extensive list would be less relevant and valuable. It would make sense to move the second task forward, restructuring the process as follows: (1) creating opportunity sequences, (2) mapping opportunities on the matrix, (3) prioritization top-5 list. Additionally, this exercise might be easier to perform in an online Miro setting, as opportunities can be more easily copied, clustered, and moved as groups.

4.4.4 User Experience Questionnaire

UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1	0,629	0,411	0,84
Perspicuity	1,583	0,673	0,440	0,6
Efficiency	1,306	0,512	0,335	-0,21
Dependability	1,028	0,491	0,321	-0,09
Stimulation	0,917	0,484	0,316	0,31
Novelty	0,361	0,761	0,497	0,56

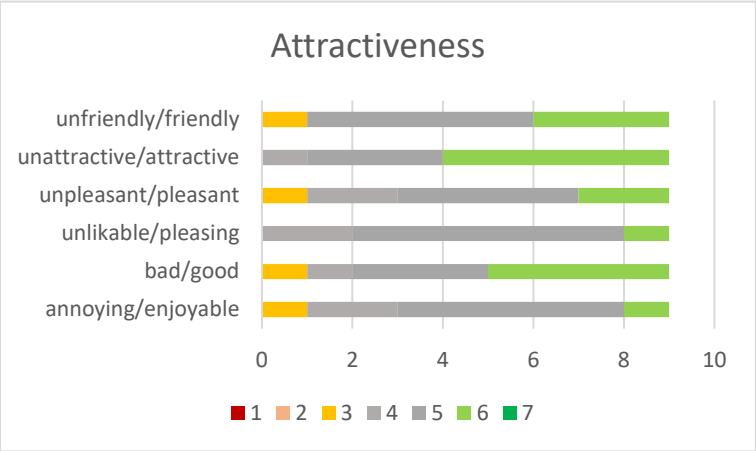
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,00	Below average	50% of results better, 25% of results worse
Perspicuity	1,58	Above Average	25% of results better, 50% of results worse
Efficiency	1,31	Above Average	25% of results better, 50% of results worse
Dependability	1,03	Below Average	50% of results better, 25% of results worse
Stimulation	0,92	Below Average	50% of results better, 25% of results worse
Novelty	0,36	Below Average	50% of results better, 25% of results worse



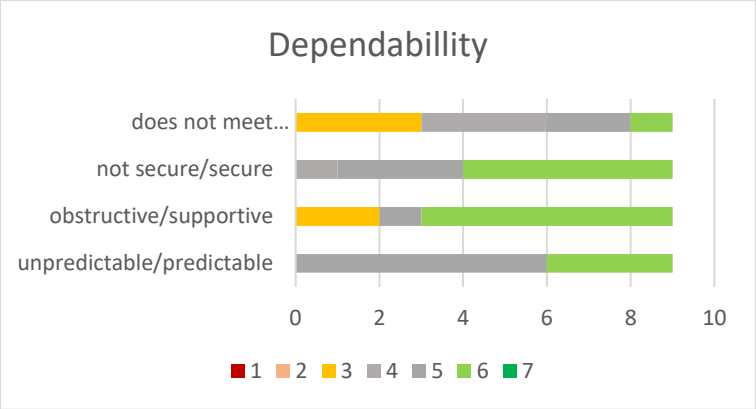
Comparing the overall user experience of the selection process shows that it scores below average on attractiveness, dependability, stimulation, and novelty. It, however, is showing above-average performance in perspicuity and efficiency. It is expected that reducing/ filtering the number of opportunities or providing more time per opportunity will positively impact the results of the dependability scale, as participants have more time to evaluate the opportunities properly. Next to this is more dominantly facilitating the session by having the facilitator read out the opportunity, potentially impacting the stimulation metric, as all participants are more actively involved in the tasks.

Moving the sequencing task forward can potentially impact the attractiveness and the efficiency of the tasks, as it prevents the idea of repetitive work. Lastly, it is likely that the nature of the participants is impactful for the metrics; if the opportunities are highly impactful on the work of the participants, they might be more stimulated and find this exercise more interesting. It is, however, hard to estimate which group will have more intrinsic motivation regarding the exercise. In the pilot, interns and supervisors are involved, who might, due to the short-term nature of their future interactions have less interest in the topic. On the other hand, the seniors that will be involved in the LS exercise might have less interest because they have other, higher priorities which, regardless of being more impactful for their long-term career, might be a low priority from a short-term perspective.

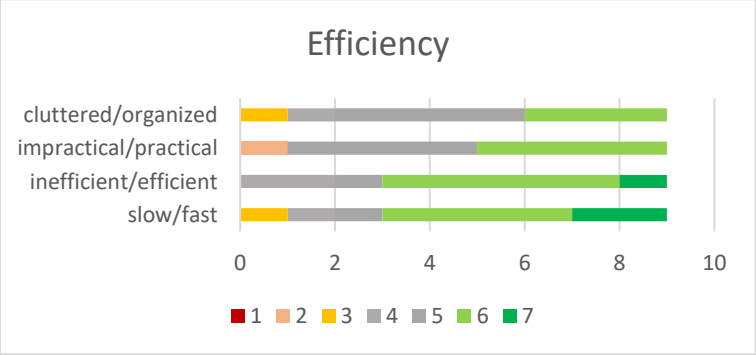
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	0,8	0,9	1,0	9	annoying	enjoyable	Attractiveness
2	1,7	1,0	1,0	9	not understandable	understandable	Perspicuity
3	0,8	1,9	1,4	9	creative	dull	Novelty
4	1,8	0,4	0,7	9	easy to learn	difficult to learn	Perspicuity
5	1,1	1,1	1,1	9	valuable	inferior	Stimulation
6	0,7	0,3	0,5	9	boring	exciting	Stimulation
7	1,0	0,5	0,7	9	not interesting	interesting	Stimulation
8	0,1	1,1	1,1	9	unpredictable	predictable	Dependability
9	1,0	1,0	1,0	9	fast	slow	Efficiency
10	0,3	1,0	1,0	9	inventive	conventional	Novelty
11	1,3	0,3	0,5	9	obstructive	supportive	Dependability
12	1,1	1,1	1,1	9	good	bad	Attractiveness
13	1,3	1,5	1,2	9	complicated	easy	Perspicuity
14	0,7	0,8	0,9	9	unlikable	pleasing	Attractiveness
15	0,2	0,9	1,0	9	usual	leading edge	Novelty
16	0,9	0,4	0,6	9	unpleasant	pleasant	Attractiveness
17	1,2	1,7	1,3	9	secure	not secure	Dependability
18	0,9	0,9	0,9	9	motivating	demotivating	Stimulation
19	1,4	0,5	0,7	9	meets expectations	does not meet expectations	Dependability
20	1,1	1,6	1,3	9	inefficient	efficient	Efficiency
21	1,6	1,0	1,0	9	clear	confusing	Perspicuity
22	1,6	1,0	1,0	9	impractical	practical	Efficiency
23	1,6	1,8	1,3	9	organized	cluttered	Efficiency
24	1,1	0,9	0,9	9	attractive	unattractive	Attractiveness
25	1,4	0,5	0,7	9	friendly	unfriendly	Attractiveness
26	0,1	1,6	1,3	9	conservative	innovative	Novelty



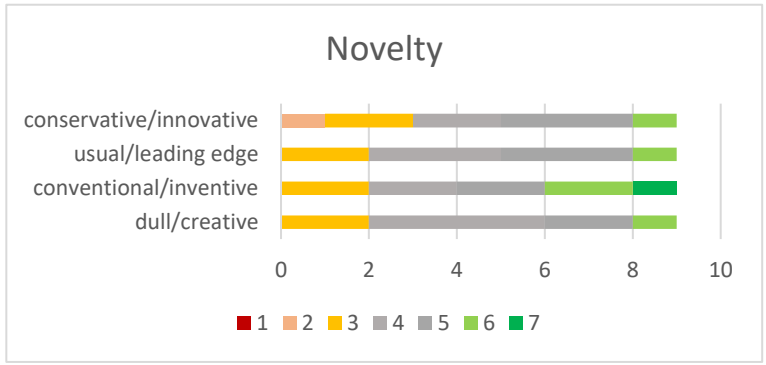
This metric might improve in the next session by providing more time and reducing the dominance of participants by taking a more active role as facilitator.



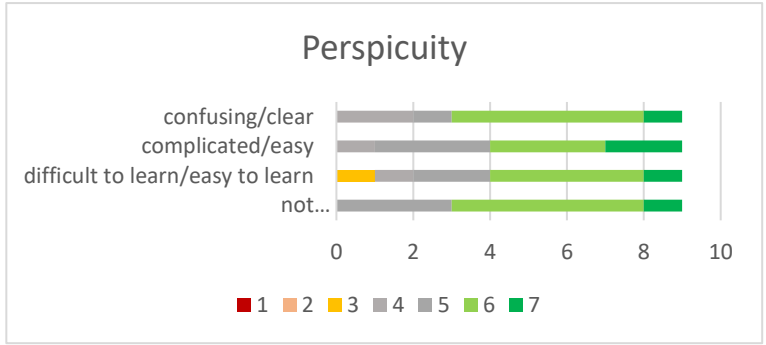
Being clear up-front about the exercise can help with managing expectations and make the process more predictable; the outcome of the selection process will, however, always be somewhat unpredictable. Additionally, the secure metric can improve when an active role of the facilitator provides a safer experience for the participants; this is likely impacted by a relatively dominant supervisor, which, due to its position, impacted the safety participants experienced during the session, potentially preventing them from speaking out.



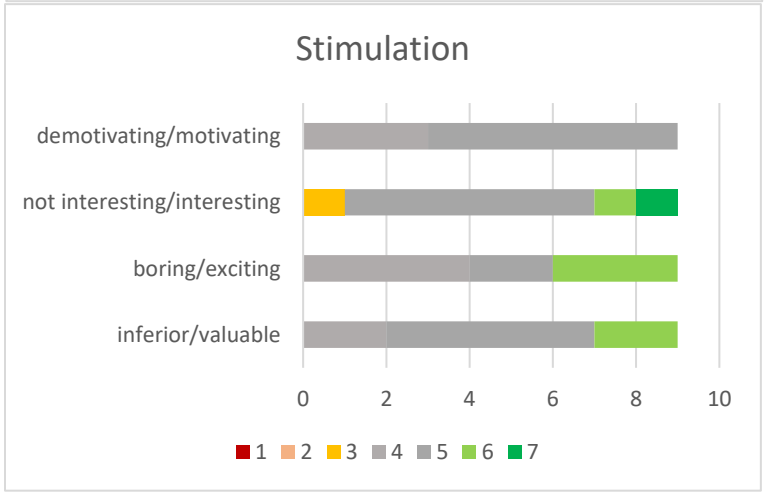
By reorganizing the structure and moving the sequencing in front of the matrix task, the process is likely more organized, practical, and efficient. The slow/fast pace of the exercise will likely be more impacted by the time that is provided per opportunity. A bit slower pace will probably positively impact the dependability of the tasks, allowing for better alignment between participants.



Since, for this task, relative conventional actions are incorporated, such as a priority/ urgency matrix, it is to be expected that the scores on novelty are not extreme. The use of sequencing and priority lists are potentially more inventive but are not differentiating enough to make it a novel exercise. Novelty is not a strong requirement, and an effective selection process is more important. Hence, no direct actions are taken in the design to improve this metric.



The exercise's perspicuity was above average, close to 'good' on the UEQ benchmark scale, and participants also expressed this. Although little to no changes are required in this field, it needs to be carefully considered that the other design changes are properly integrated and communicated to the participants.



For many participants, the exercise was a bit boring and uninteresting, showing mediocre intrinsic motivation. This might be because the target group is interns and supervisors whose primary task is not directly improving the intern experience for future interns. If, based on these metrics, participants were indeed not strongly motivated, this might also have impacted the rest of the metrics, as enthusiasm for the topic can broadly impact various metrics.

It is expected that senior professionals who attempt to improve internal operations daily might be more appreciative of the identified opportunities as they can support the creation of strategies that have a long-term impact. Potentially, however, these seniors, who are always facing temporal challenges, might have short-term priorities that overshadow the session, reducing their intrinsic motivation as well, which is a risk that is hard to influence or estimate.

4.5 Opportunity Modelling

The Participatory Modelling Session is adopted from the System Mapping Academy and hasn't been adjusted to firstly test its performance during the pilot. For the process, a time indication was given for each step; the process consisted of 5 main steps, each of which is home to 3-to 5 sub-tasks. Each step was appointed 15 minutes in total except for step 3, for which 40 minutes were estimated. In total, the workshop was scheduled within a timeslot of two hours. Providing 20 minutes of undedicated time for breaks, introductions, explanations, and delays. The time indications appeared to be quite accurate, and the two-hour slot was about enough to go through the entire process. Only the final sub-step was skipped, as time was running short, and its relevance was too low.

Overall, the workshop seemed to be really well received, and the sessions seemed to be way more effective when compared to the benchmark. The benchmark session attempted the same thing but failed mostly during this particular step. The participants found it too complex and difficult to do. The Participatory modeling methods from the System Mapping Academy are generally highly comparable to the original process. The big difference, however, is that each step is broken down into more small steps, which are all very well explained, making it very easy for participants to follow the process. Next to this the workshop offers an example, extensive facilitator guidance, and background information.

4.5.1 End of Survey Comments

Do you have any tips?	**Do you have any tops?**	**Other feedback?**
I think there were points where you could have hurried us along, but this is in hindsight.	Good facilitation, clearly explained when we had questions, I think it was productive	Im not sure what the next steps are but im keen to talk about solutions already
Good workshop	Really good organized, good guidance throughout	Maybe guide the discussion a bit more if we roll of a bit too much from the topic
Perhaps prioritizing the variables before mapping would create a smoother mapping, as it was quite chaotic now	Good diverge converge structure	
What I already said: I get the point, but at some point the exercise gets more complicated than the problem.	I think you prepared the session very well and hosted it very well. All in all it looked professional and came across professional.	If you want me to explain all this, I'm fine with a call. I get that this might be a lot of info and maybe not what you expected. So what you want :)

Based on the comments, the facilitation and organization were considered to be good and professional. However, the workshop might have benefited from more strong facilitation guidance.

A more process-driven suggestion was that variables could be prioritized before being made. This would make it easier to place the key variables in the model first before populating it with more secondary variables. This way, task three might become easier and less mentally demanding. Since task 3 was also considered slightly more demanding and difficult in the data, it is worth integrating this into the workshop to improve the flow of the session.

4.5.2 NASA Task Load Index

4.5.2.1 P-TLX – Opportunity Modelling – Pilot vs Bench

TLX Metric	Mean P-MDL	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,0	6,8	-2,8	1,0	0,053	0,875
Performance	7,5	7,3	0,3	9,0	0,881	-0,125
Mental Demand	5,0	7,0	-2,0	2,5	0,137	0,688
Temporal Demand	2,8	6,3	-3,5	1,5	0,078	0,813
Frustration	3,0	1,0	2,0	13,0	0,134	-0,625

4.5.3 Task Specific Difficulty

4.5.3.1 Task-specific feedback

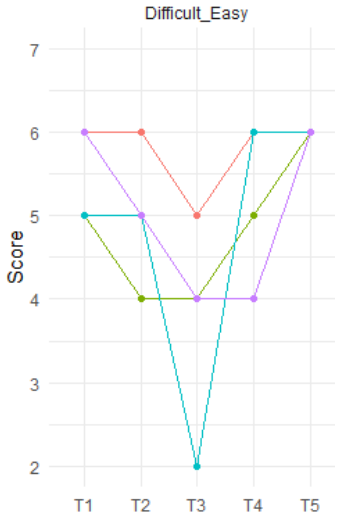
T1	T2	T3	T4	T5
na	na	na	na	na
.	is quite mentally demanding because of the brainstorming	,	.	.
5	Not rushed	Hard to make structure from chaos, unsure about accuracy of results but dont know better.	x	x
No	no	I think there were too many factors and with everything being somewhat connected to everything else, it quickly became very complicated. Also since we are working towards 'getting better onboarding', which isn't the most complicated problem in the world, this somewhat seems like overkill. But still I get the point of the session. I like it, I like describing the problem well and getting to an answer, but as a very 'beta / right or wrong' kind of person, these kind of broad maps start to seem like 'just a map' and not the actual truth, because there is so much discussion about it.	To continue on my last remark: At this point I start feeling like this exercise starts surpassing the purpose. We could have gotten to answers to our problems 2 steps before this. I think the entire map so big now that it can't be summarized by 'lack of communication', it covers way more than this.	No

Based on the comments on the tasks, some considered the topic that was selected for the exercise to be overkill, stating that the exercise might become more complex than the problem. This might have contributed to some frustration scores. Additionally, the accuracy of the mapping was questioned as the participants were constantly discussing how variables interacted. Additionally, according to one participant, it might have been more effective to target solutions more directly, which might have resulted in effective solutions at an earlier moment in this process. (Part of the overall conclusion).

One insight from this workshop is that during the selection, it should be clear what the intention is of the follow-up workshop and that the challenge that should be selected is suitable for the modeling exercise. Thus, the selection exercise should not directly select the top priority from the list, as this opportunity might be a highly important but easy-to-achieve goal. Instead, it should target the most complex but highly potential opportunity, as this is where the modeling workshop is likely to excel in providing value.

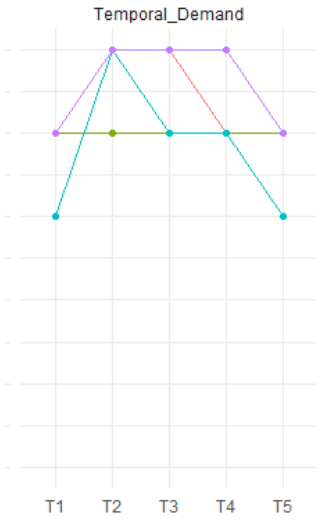
4.5.3.2 Internal Task Comparison

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	T1	T2	5,50	5,00	0,50	0,43	-0,38
Difficulty	T1	T3	5,50	3,75	1,75	0,05	-0,88
Difficulty	T1	T4	5,50	5,25	0,25	0,87	-0,13
Difficulty	T1	T5	5,50	6,00	-0,50	0,18	0,50
Difficulty	T2	T3	5,00	3,75	1,25	0,17	-0,63
Difficulty	T2	T4	5,00	5,25	-0,25	0,76	0,19
Difficulty	T2	T5	5,00	6,00	-1,00	0,07	0,75
Difficulty	T3	T4	3,75	5,25	-1,50	0,13	0,69
Difficulty	T3	T5	3,75	6,00	-2,25	0,02	1,00
Difficulty	T4	T5	5,25	6,00	-0,75	0,19	0,50

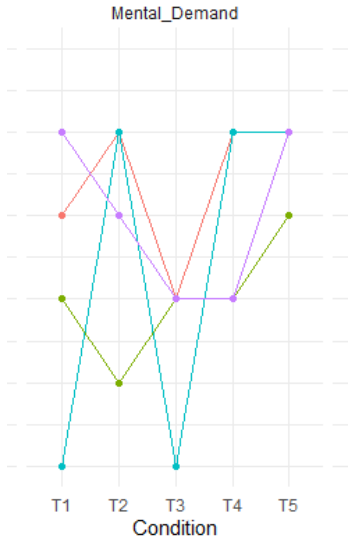


Comparing the difficulty of the tasks inside the workshop shows that there are no significant differences between them. The strongest differences are found between exercise 1 vs. 3 and 3 vs. 5, which, with a P value of 0,09, are closest to the 0,05 value for significance.

Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Temporal Demand	T1	T2	5,50	6,75	-1,25	0,05	0,88
Temporal Demand	T1	T3	5,50	6,50	-1,00	0,09	0,75
Temporal Demand	T1	T4	5,50	6,25	-0,75	0,13	0,63
Temporal Demand	T1	T5	5,50	5,75	-0,25	0,61	0,25
Temporal Demand	T2	T3	6,75	6,50	0,25	0,61	-0,25
Temporal Demand	T2	T4	6,75	6,25	0,50	0,25	-0,50
Temporal Demand	T2	T5	6,75	5,75	1,00	0,06	-0,81
Temporal Demand	T3	T4	6,50	6,25	0,25	0,61	-0,25
Temporal Demand	T3	T5	6,50	5,75	0,75	0,13	-0,63
Temporal Demand	T4	T5	6,25	5,75	0,50	0,26	-0,44



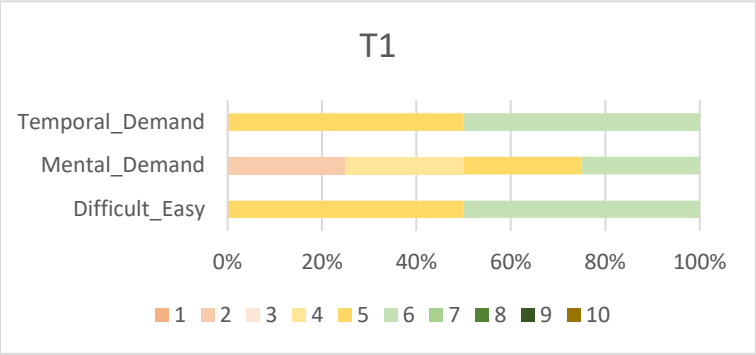
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Mental Demand	T1	T2	4,25	5,00	-0,75	0,55	0,31
Mental Demand	T1	T3	4,25	3,50	0,75	0,44	-0,38
Mental Demand	T1	T4	4,25	5,00	-0,75	0,65	0,25
Mental Demand	T1	T5	4,25	5,75	-1,50	0,16	0,63
Mental Demand	T2	T3	5,00	3,50	1,50	1,00	-0,63
Mental Demand	T2	T4	5,00	5,00	0,00	1,00	0,00
Mental Demand	T2	T5	5,00	5,75	-0,75	0,50	0,31
Mental Demand	T3	T4	3,50	5,00	-1,50	0,13	0,63
Mental Demand	T3	T5	3,50	5,75	-2,25	0,02	1,00
Mental Demand	T4	T5	5,00	5,75	-0,75	0,40	0,38



Between the exercises within the Modeling workshop, no significant differences were found concerning difficulty, temporal demand, and mental demand. Especially the temporal demand seems to score consistently across all 5 exercises, meaning that the time indication per exercise is properly distributed. A slight drop in difficulty and temporal demand is visible when checking the third task; this exercise was the creation of the actual model based on the defined variables in task 2. The main reason for this is that besides the roughly 20 variables that were identified, the underlying connection needed to be assessed and understood, making this a relatively mentally demanding and complex task. Considering the complexity of the task, and experiences with a previous attempt to achieve modeling, the exercise is considered quite successful in supporting the participants in performing the task.

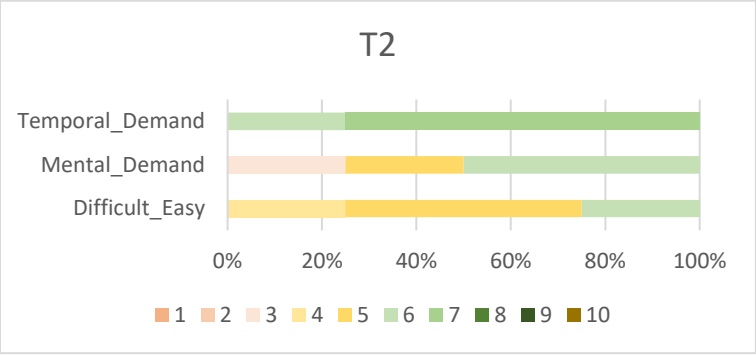
4.5.3.3 Task Specific Data Distribution

4.5.3.3.1 P-TS – Opportunity Modelling – Task 1



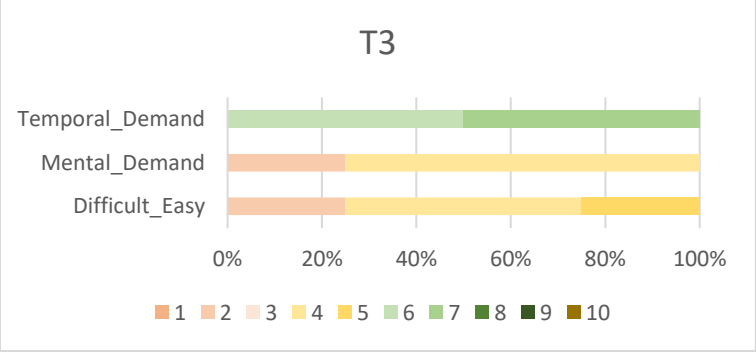
The first task was about framing the opportunity using a framing template, asking to specify a specific topic, problem description, research question, goals, and scope that shows what should be in and out of focus during the exercise. The diagram shows one participant who found this exercise quite mentally demanding (2 out of 10), but the others found this to be relatively (4,5 & 6 out of 10). Both difficulty and temporal demand seem to score comparable, being either 5 or 6 out of 10, showing that the exercise was experienced as an average difficult and rushed exercise.

4.5.3.3.2 P-TS – Opportunity Modelling – Task 2

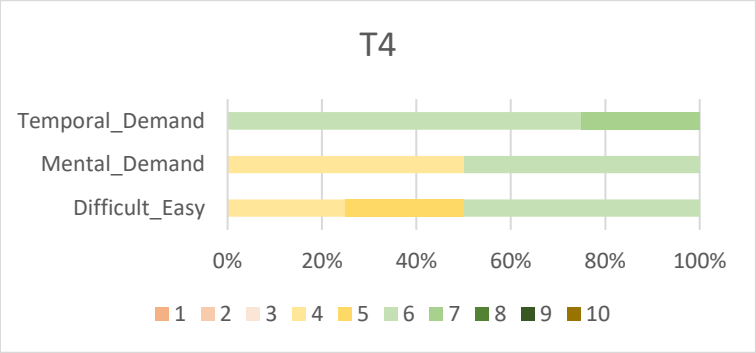


The second exercise asked the participant to identify the main variable and secondary variables surrounding this main variable. This exercise scored slightly better in both temporal and mental demand but was experienced as a little more difficult than the first exercise.

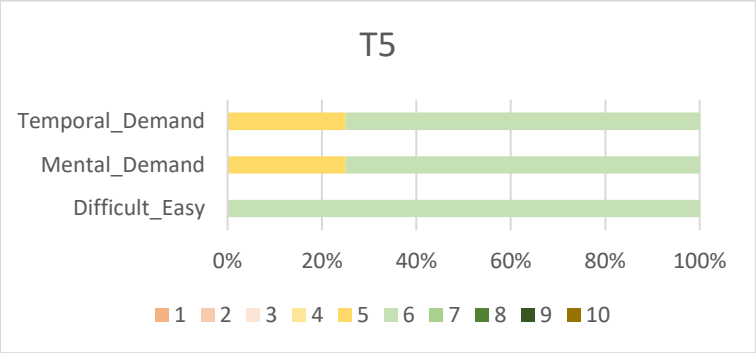
4.5.3.3.3 P-TS – Opportunity Modelling – Task 3



4.5.3.3.4 P-TS – Opportunity Modelling – Task 4



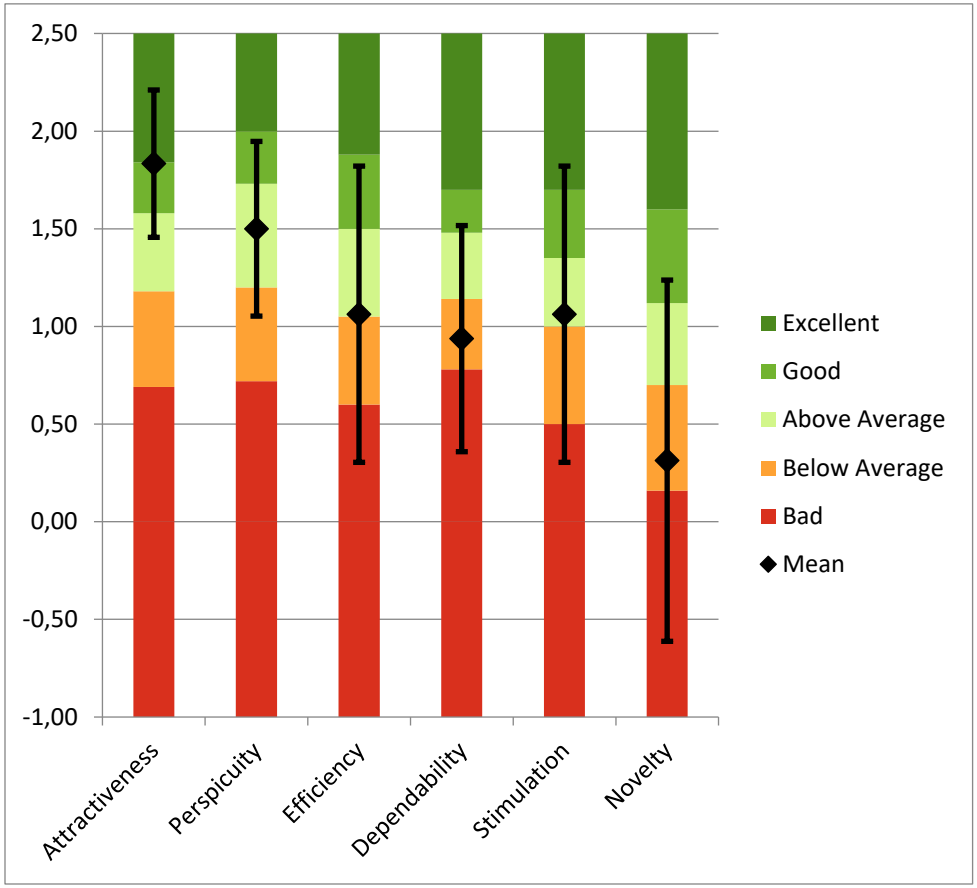
4.5.3.3.5 P-TS – Opportunity Modelling – Task 5



4.5.4 User Experience Questionnaire

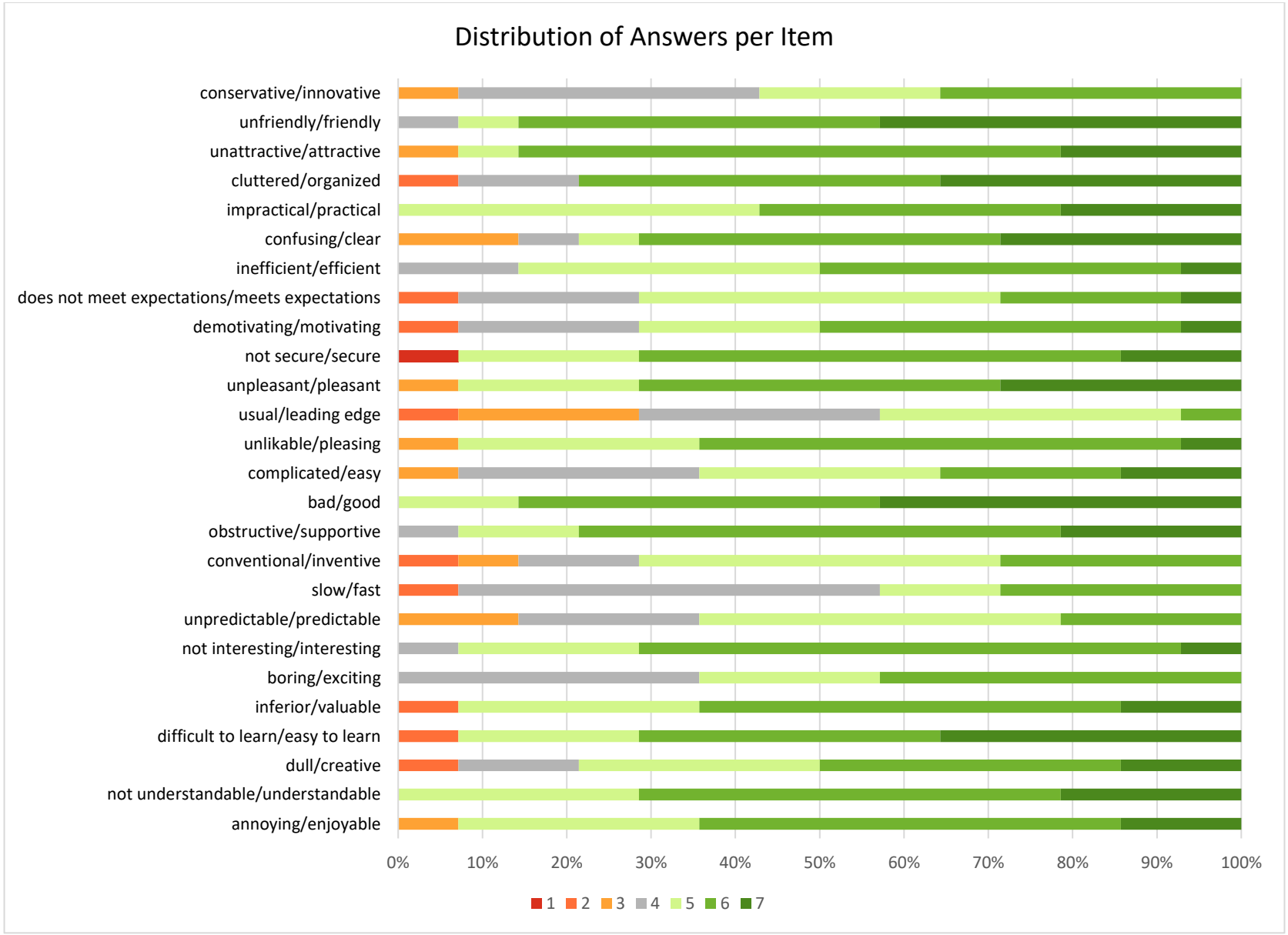
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,833	0,15	0,377	0,66
Perspicuity	1,5	0,21	0,447	0,75
Efficiency	1,063	0,6	0,758	0,48
Dependability	0,938	0,35	0,579	0,55
Stimulation	1,063	0,6	0,758	0,84
Novelty	0,313	0,89	0,925	0,87

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,83	**Good**	10% of results better, 75% of results worse
Perspicuity	1,50	**Above Average**	25% of results better, 50% of results worse
Efficiency	1,06	**Above Average**	25% of results better, 50% of results worse
Dependability	0,94	**Below Average**	50% of results better, 25% of results worse
Stimulation	1,06	**Above Average**	25% of results better, 50% of results worse
Novelty	0,31	**Below Average**	50% of results better, 25% of results worse



Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,8	0,9	1,0	4	annoying	enjoyable	Attractiveness
2	1,8	0,3	0,5	4	not understandable	understandable	Perspicuity
3	0,8	2,3	1,5	4	creative	dull	Novelty
4	1,8	0,3	0,5	4	easy to learn	difficult to learn	Perspicuity
5	1,3	0,9	1,0	4	valuable	inferior	Stimulation
6	0,8	1,6	1,3	4	boring	exciting	Stimulation
7	1,0	0,7	0,8	4	not interesting	interesting	Stimulation
8	-0,3	0,9	1,0	4	unpredictable	predictable	Dependability
9	1,3	0,9	1,0	4	fast	slow	Efficiency
10	0,3	1,6	1,3	4	inventive	conventional	Novelty
11	1,0	0,7	0,8	4	obstructive	supportive	Dependability
12	1,8	0,3	0,5	4	good	bad	Attractiveness
13	1,3	0,3	0,5	4	complicated	easy	Perspicuity
14	1,5	0,3	0,6	4	unlikable	pleasing	Attractiveness
15	0,0	0,7	0,8	4	usual	leading edge	Novelty
16	1,8	0,3	0,5	4	unpleasant	pleasant	Attractiveness
17	1,8	0,3	0,5	4	secure	not secure	Dependability
18	1,3	0,3	0,5	4	motivating	demotivating	Stimulation
19	1,3	0,9	1,0	4	meets expectations	does not meet expectations	Dependability
20	0,8	1,6	1,3	4	inefficient	efficient	Efficiency
21	1,3	0,9	1,0	4	clear	confusing	Perspicuity
22	1,0	2,0	1,4	4	impractical	practical	Efficiency
23	1,3	0,9	1,0	4	organized	cluttered	Efficiency
24	1,8	0,3	0,5	4	attractive	unattractive	Attractiveness
25	2,5	0,3	0,6	4	friendly	unfriendly	Attractiveness
26	0,3	0,9	1,0	4	conservative	innovative	Novelty


Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0 1 2 3 4 5
1 2 3 4 5 6 7
Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0 1 2 3 4 5
1 2 3 4 5 6 7
Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0 1 2 3 4 5
1 2 3 4 5 6 7
Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0 1 2 3 4 5
1 2 3 4 5 6 7
Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0 1 2 3 4 5
1 2 3 4 5 6 7
Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0 1 2 3 4 5
1 2 3 4 5 6 7


Distribution of Answers per Item
conservative/innovative
unfriendly/friendly
unattractive/attractive
cluttered/organized
impractical/practical
confusing/clear
inefficient/efficient
does not meet expectations/meets expectations
demotivating/motivating
not secure/secure
unpleasant/pleasant
usual/leading edge
unlikable/pleasing
complicated/easy
bad/good
obstructive/supportive
conventional/inventive
slow/fast
unpredictable/predictable
not interesting/interesting
boring/exciting
inferior/valuable
difficult to learn/easy to learn
dull/creative
not understandable/understandable
annoying/enjoyable
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
1
2
3
4
5
6
7

4.6 Pilot Design Changes

4.6.1 Research protocol

- Clarify the direction of all scales (might be confusing in the TLX)
- Highlight elements in the reading part to support scanning over it.

4.6.2 Context Capturing

Perspicuity improvements

- Overall, improve/ implement more explainers (workshop should be almost fully autonomous)
- Explain the link between BCM & CJM Elements
- Perspicuity should improve, as well as below average vs benchmark and participant feedback.
- PAINS & GAINS Should be each other's opposites, a very clear definition is needed
- Clarify - Keep journey phases broad, details can be in the JTBD
- Clarify - Steps in the BCM (Partners vs Customer) give examples
- Explain the underlying goal of the session, which helps focus and drive the participant!
- Use a good title or research question to focus the exercise!

Time balancing

- More Time for CJM

Allow Preparation

- Introductory message to allow participants to prepare for the interview.

Highlight respect for others' input.

- Checking each other's work in combination with hierarchy differences might cause feelings of unease. Implement of a good explanation of the purpose in the introduction to mitigate this risk.

Facilitator guidance

- Facilitator guide (How to nudge, push & support participants?)
- Provide Revision questions/ prompts

Other

- Always Incorporate familiarity check with Miro
- Pre-tag post-its
- 2-4 words per post-it!

4.6.3 Opportunity Identification

Simplify task 3 -

A difference can be found between Task 3 (formulating new strategies) & Task 4 (selection of strategies). The difference is found both in the difficulty and in the temporal demand. However, not with strong significance, it has to be noted that task 3 is the most difficult and mentally demanding of all tasks, whereas the selection exercise is the easiest and least mentally demanding one. It is nice that these are following each other up, giving participants some time to recover. It could be considered to simplify task 3, but as of yet, no indication of how to improve it has been identified, and since the difference is quite low, it also might not be a major priority.

Perspicuity

Both the interview and the OI workshop perform low on perspicuity in comparison to the benchmark. Part of this might be due to the fact that the benchmark products perform higher than a workshop, which people are not familiar with; however, based on participant feedback, I suspect that the perspicuity can be improved by integrating the design ideas and improving the explanations.

Timing

The first tasks require more time as participants familiarize themselves with the blueprint. The same is true for the second task, but to a lesser extent.

4.6.4 Opportunity Selection

For the OS session various change were made. Firstly, the format was shifted from physical to digital for logistic reasons. Secondly the Matrix Mapping and Grouping/Sequencing exercise were switched out to improve the sessions flow. Lastly a template was designed for the Miro board in which more was explained to the participants in writing.

4.6.5 Opportunity Modelling

For the Opportunity Modeling, minor changes were made; one element was added, which was about selecting the key variables before starting to model. Additionally, the timing slightly changed to balance temporal load.

5. Use case - Data analysis

Results from Data Analysis

CC-Task Load

Task load - Starting Participants

Comparing the Starting Participants (original method) to the benchmark shows a significant decrease in Effort (P=0,019), Performance (P=0,017), Mental Demand, (P=0,019) and Temporal Demand (P=0,017) (Table 1).

TLX Metric	Starting Participants	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	3,0	7,5	-4,5	0,0	0,019	1,000
Performance	2,6	7,0	-4,4	0,0	0,017	1,000
Mental	3,6	8,0	-4,4	0,0	0,019	1,000
Temporal	3,4	6,3	-2,9	0,0	0,017	1,000
Frustration	1,6	1,3	0,4	11,5	0,794	-0,150

Table 1 - Task Load statistics (Starting participants vs Benchmark)

Task load - Follow-up Participants

For the Follow-Up Participants (new method) a significant difference is found for Effort (P=0,05), Performance (P=0,014) and Mental Demand (P=0,013), Temporal Demand is nearing significance (P=0,083) (Table 2).

TLX Metric	Follow-Up Participants	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,6	7,5	-2,9	2,0	0,050	0,722
Performance	3,1	7,0	-3,9	2,0	0,014	0,889
Mental	4,7	8,0	-3,3	5,0	0,013	0,889
Temporal	4,6	6,3	-1,7	6,5	0,083	0,639
Frustration	2,4	1,3	1,2	21,5	0,636	-0,194

Table 2 - Task Load statistics (Follow-up participants vs Benchmark)

Task load - Starting vs. Follow-up Participants

Comparing First and Follow-Up participants did not result in any significant differences. However, it is notable that all indicators scored slightly lower on TL for the Starting Participants. Frustration reaches no significance for any of the three comparisons (Table 3) (Appendix 5.1.2).

TLX Metric	Starting Participants	Follow-Up Participants	Mean Difference	W	P-Value	Effect Size
Effort	3,0	4,6	-1,6	14,0	0,27	0,38
Performance	2,6	3,1	-0,5	20,5	0,83	0,09
Mental	3,6	4,7	-1,7	18,0	0,58	0,20
Temporal	3,4	4,6	-1,2	13,5	0,24	0,40
Frustration	1,6	2,4	-0,8	21,0	0,89	0,07

Table 3 - Task Load statistics (Starting vs Follow-up participants)

CC-User Experience

UX - Starting Participants

The Starting Participants filled in empty templates and rated the UX consistently higher than the Follow-Up participants. Attractiveness and Perspicuity both had excellent scores for this group (90%>), scoring high on all underlying sub-scale items. Efficiency was deemed good (<90%), as it was seen as very well organized. Dependability was above average (<75%), scoring a bit lower on predictability and meeting expectations. Stimulation had a good score (<90%), as the session was experienced as highly valuable, interesting, and motivating. The Novelty score was above average (<75%), reflecting that the session was highly innovative and quite creative but not very leading edge or inventive (Appendix 5.1.4.1.1).

UX - Follow-Up Participants

The Follow-Up participants reviewed and revised the work of their predecessors. Attractiveness was good (<90%), scoring well on all underlying items. Perspicuity was above-average (<75%), scoring high on understandability and ease of learning, but had lower scores on the clarity and complexity items. Efficiency had an above-average score (<75%), scoring highly on practicality and efficiency but low on speed, and it was considered a bit cluttered by some. Dependability was also above average (<75%), the session was found very supportive but also a bit unpredictable. Stimulation reported an above-average score, being very interesting and valuable but, for some, not too exciting. On Novelty, the session scored below average (<50%), scoring low on innovativeness, as participants found the session quite usual and conventional (Appendix 5.1.4.1.2).

OI-Task Load

Significant Reduction in Effort and Temporal Demand

Compared to the benchmark (Table 4), the OI-session did significantly reduce its TL on Effort (P=0,016) and Temporal demand (P=0,038) and closely nears significance on Mental demand (P=0,085). Interestingly, the Perceived Performance score showed a lower decrease than the other indicators (Mean Difference=1.2). The increase in Frustration is closely nearing significance (P=0,053). (Appendix 5.2.2)

TLX Metric	Mean OI-Session	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,8	6,8	-1,9	4,0	0,016	0,818
Performance	6,1	7,3	-1,2	6,0	0,138	0,523
Mental	5,0	7,0	-2,0	37,0	0,085	0,614
Temporal	3,5	6,3	-2,8	8,5	0,038	0,727
Frustration	4,0	1,0	3,0	10,5	0,053	-0,682

Table 4 - Task Load statistics Opportunity Identification

OI-User Experience

Low Perspicuity, Efficiency and Dependability

The OI-session scored on Attractiveness above average (<75%). It scored badly (<25%) on the Perspicuity UEQ-benchmark, as various participants found the session confusing and complicated. On Efficiency, the score was below average (<50%), as the session could have been a bit faster and was also found to be somewhat impractical. Dependability also scored below average (<50%), as it was seen as unpredictable and, for some, did not meet expectations. Stimulation was good (<90%), rating highly interesting, motivating, and valuable. Novelty was above average (<75%), scoring highly on innovativeness and leading edge; opinions were strongly divided on the dull/creative scale (Appendix 5.2.4).

OS-Task Load

Comparable Task Load to OI-Session

The OS-session scored consistently on the TLX-scale (Table 5), ranging between means of 4 and 5. The Effort (P=0,08), Performance (P=0,137), and Mental demand (P=0,028) all decreased in comparison to the benchmark, with noticeable effect size (>0,6). However, only the Mental demand achieved significance, due to a small sample size (n=4). The effects on the temporal demand were, in comparison, small (Effect size=0,313). The increase in Frustration is not significant, regardless of having the largest mean difference. Overall, the TL shows an equal decrease on perceived performance and Effort and Mental Demand, Temporal Demand is slightly less decreased, and Frustration shows quite a big but not significant increase. (Appendix 5.3.2)

TLX Metric	Mean OS-Session	Mean Bench-mark	Mean Difference	W	P-Value	Effect Size
Effort	4,5	6,8	-2,3	0,0	0,080	0,813
Performance	5,0	7,3	-2,3	1,5	0,137	0,688
Mental	4,5	7,0	-2,5	2,5	0,028	1,000
Temporal	5,0	6,3	-1,3	12,5	0,557	0,313
Frustration	4,0	1,3	2,7	5,5	0,234	-0,563

Table 5 - Task Load statistics Opportunity Selection

OS-User Experience

Perspicuity & Efficiency Pilot vs Use-case

The UX of the OS-session is the lowest scoring across all sessions for each metric. Attractiveness was below average (<50%) as the session was experienced as a bit unpleasant. Perspicuity scored bad (<25%), as it was complicated, not understandable, and confusing. Efficiency also scored bad (<25%), receiving low scores on practicality, efficiency, and speed. Again, Dependability scored bad (<25%), as the session was found to be obstructive and not meeting expectations. Stimulation was below average (<50%), being mostly scored as boring. Novelty also was below average (<50%) as it was perceived as dull, conventional, and usual (Appendix 5.3.4).

OM-Task Load

The duration of this session was cut in half due to an unexpected development, resulting in all participants being unable to participate for the full session. Thus, it was decided to only perform the first three phases, reducing time intended per task and requiring strict timekeeping by the facilitator. Due to the time constraints of this study, no follow-up session could be planned to test the last two tasks (Reflect & Leverage).

High Mental & Temporal Demand

The Mental (P=0,301) and Temporal Demand (P=0,372) did reduce but were relatively low and did not near the significance threshold. Effort was, however, significantly reduced (P=0,026), as well as perceived performance, which nears significance (P=0,078). Frustration did increase but reached no significance (P=0,457) (Table 6). (Appendix 5.4.2)

TLX Metric	Mean OM-Session	Mean Bench-mark	Mean Difference	W	P-Value	Effect Size
Effort	4,0	6,8	-2,8	-2.75	0,026	1,000
Performance	5,0	7,3	-2,3	-2.25	0,078	0,813
Mental	6,0	7,0	-1,0	-1.00	0,301	0,500
Temporal	5,0	6,3	-1,3	-1.25	0,372	0,438
Frustration	2,8	1,3	1,5	1.75	0,457	-0,375

Table 6 - Task Load statistics Opportunity Modelling

OM-User Experience

Low Perspicuity, Dependability, and Efficiency
Perspicuity scored badly, (<25%) as it was found confusing, complicated, and difficult, not very understandable. Also, Dependability scored badly (<25%), primarily due to a low score on predictability. Efficiency scored below average (<50%), as some found it a bit slow and inefficient. Stimulation had very good scores (>75%) and was found interesting and highly valuable. Novelty also scored good (>75%) as the session was found to be creative, leading edge, and innovative (Appendix 5.4.4).

5.1 Context Capturing

5.1.1 End of Survey Comments

5.1.1.1 Starting Participants

Do you have any tips?	**Do you have any tops?**	**Other feedback?**
maybe specify that 'S&C' is 'S&C Life Sciences', as H&PS is also part of S&C along with many other teams	very well-structured Miro Board that was easy to follow and complete	
none	none	none
.	.	.
not at the moment	great that the presenter thinks along	it would be nice to understand how exactly this information is going to be incorporated into next steps
No	No	Goodluck i guess!

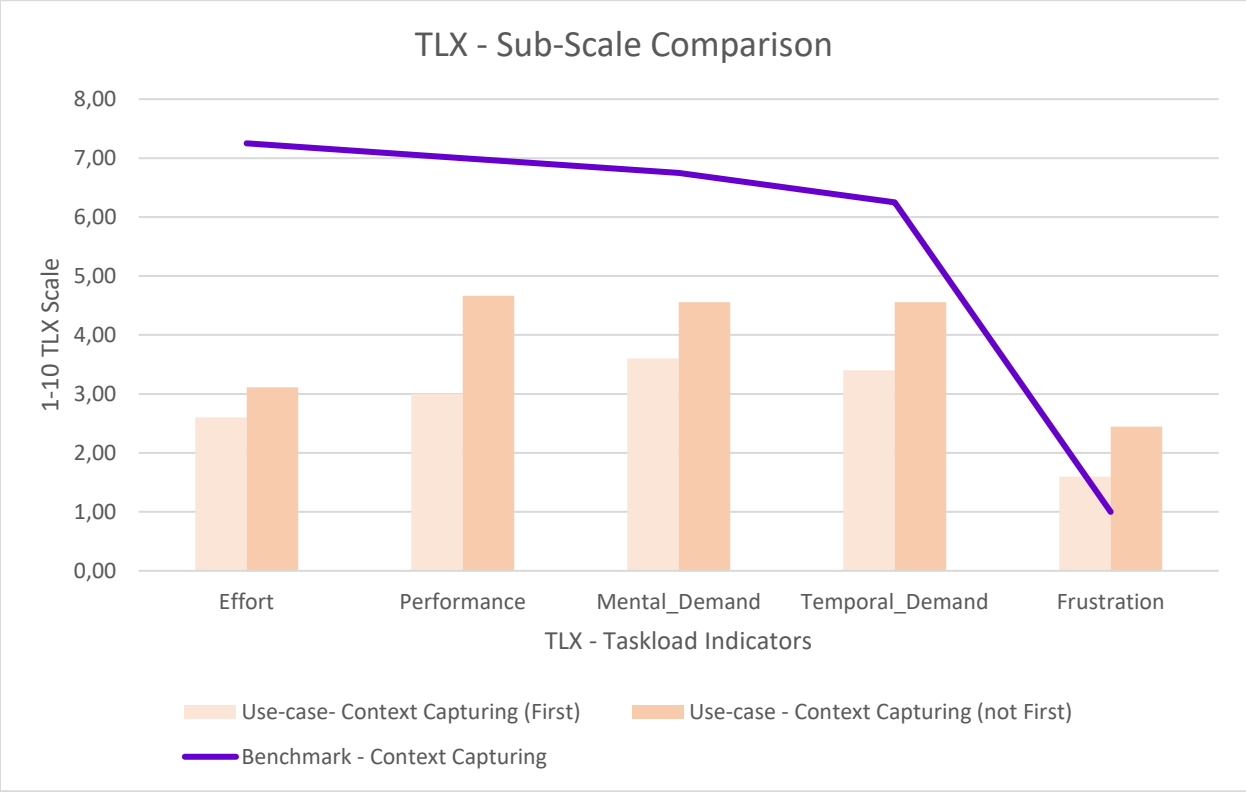
5.1.1.2 Follow-Up Participants

Do you have any tips?	**Do you have any tops?**	**Other feedback?**
Explain a bit more why you do it, what you research and what you want to get out of it and how it could support the practice.	Professional approach with Miro board and questionaire.	A bit more time for scheduling an interview
N/A	N/A	N/A
Official definition of other services should be giving. S&C?	No	x
Benefits of canvas is to discuss it in the group. Doing one on one on forehand helps us (ACN LT) to prepare but also missing the interaction	Support and prep was done great	Feels a bit overwhelming to do a canvas instead of an interview (which is mentioned in the invite)
More emphasis on LS as answers started becoming Accenture-generic	Board could have been simplified to reduce duplication	x
It could be benefitial to share miro board before the session for preparation	It could be better to have this session within a group	x
None	Good to show input from others, can be build upon	x
maybe have a break in between	cam on is definetely a plus	enjoyable experience!
for this session of an hr it's important to have the prepwork done, as my colleague has done. Otherwise it wasn't possible to do in an hr. I'm also used to the	Miro was well prepared. The timeblocks worked fine for me. As mentioned, I get the purpose and the effect of it. It was great to build on the input already provided.	thanks, looking forward to the output of 28th of March - the devil is in the action afterwards. So how to make it tangible, test it, etc. It would be great to end with more

formats/methods which helped as well. For 'novice' in this way of working you need more time and explanation. 2. it could help to hide the content and show it when you progress during the meeting, its overwhelming if you see it all. 3. the prep upfront was a bit short notice. 4. the goal is to find synergy options - it's crucial to understand what other partners in the network 'do' and don't to find the connections / opportunities. It still felt as an 'internal focus' excercise when I look back to the journey and Value canvas - the area we're looking for is the combination / touchpoints and the value props in that area. The purpose of the journey wasn't really clear to me > this will give insight in how we can improve our internal sales process, however, this is already 'known', and again, not focused on the synergy. So I hope this will be the focus on 28th of March?		than a summary and report full of 'advice'. Not sure what the plans are, but try to make sure attendees have the time (and 10 reminders) to prep before we meet IRL...

5.1.2 NASA Task Load Index

5.2.1 Use-Case – Context Capturing – TLX - Summary



Comparing the individual items in the TLX between the benchmark and the Context Capturing Sessions shows that for both the first and not first participants, the Effort, Mental Demand, and Temporal demand significantly decreased. Additionally, the Perceived Performance of the participants decreased as well. A small and consistent higher score is visible between the first and not first participants; this difference, however, does need to meet the significance threshold.

Since the design change to this session only impacted the not first participants, it is surprising that the first participants experienced an even lower task load. It is not definitive as to why this occurred; it might be due to perspicuity improvements, time balancing, and other minor changes that were implemented based on the Pilot. Additionally, the difference in context differed somewhat, targeting only the Life Sciences

domain instead of the total service spectrum per group. Lastly, the sample is different, including different participants for ID & IX and new groups, like ESP, H&PS, and S&C.

In **conclusion**, it is impossible to conclude that integrating the modified Delphi method had any positive effect on the task load. The slight difference between first and not first participants does indicate that the design change had an overall negative effect on the task load. However, this cannot conclusively be stated, as the effect and sample size were not large enough to reach the significance threshold.

5.1.2.2 LS-TLX – Interviews – First vs Bench

TLX Metric	Starting Participants	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	3,0	7,5	-4,5	0,0	0,019	1,000
Performance	2,6	7,0	-4,4	0,0	0,017	1,000
Mental Demand	3,6	8,0	-4,4	0,0	0,019	1,000
Temporal Demand	3,4	6,3	-2,9	0,0	0,017	1,000
Frustration	1,6	1,3	0,4	11,5	0,794	-0,150

5.1.2.3 LS-TLX – Interviews – Not First vs Bench

TLX Metric	Follow-Up Participants	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,6	7,5	-2,9	2,0	0,050	0,722
Performance	3,1	7,0	-3,9	2,0	0,014	0,889
Mental Demand	4,7	8,0	-3,3	5,0	0,013	0,889
Temporal Demand	4,6	6,3	-1,7	6,5	0,083	0,639
Frustration	2,4	1,3	1,2	21,5	0,636	-0,194

5.1.2.4 LS-TLX – Interviews – Total vs Bench

TLX Metric	Starting Participants	Follow-Up Participants	Mean Difference	W	P-Value	Effect Size
Effort	3,0	4,6	-1,6	14,0	0,27	0,38
Performance	2,6	3,1	-0,5	20,5	0,83	0,09
Mental Demand	3,6	4,7	-1,7	18,0	0,58	0,20
Temporal Demand	3,4	4,6	-1,2	13,5	0,24	0,40
Frustration	1,6	2,4	-0,8	21,0	0,89	0,07

5.1.2.5 LS-TLX – Interviews – First LS vs First Pilot

TLX Metric	LF-f	Pilot-f	Mean Difference	W	P-Value	Effect Size
Effort	3,00	7,5	-4,5	5,0	0,077	0,667
Performance	2,60	7,0	-4,4	0,0	0,007	1,000
Mental Demand	3,60	8,0	-4,4	6,5	0,136	0,567
Temporal Demand	3,40	6,3	-2,9	8,5	0,258	0,433
Frustration	1,60	1,3	0,3	19,0	0,508	-0,267

5.1.2.6 LS-TLX – Interviews – Not First LS vs Not First Pilot

TLX Metric	LS-nf	Pilot-nf	Mean Difference	W	P-Value	Effect Size
Effort	4,56	6,8	-2,2	25,5	0,267	-0,417
Performance	3,11	3,3	-0,1	2,0	0,014	0,889
Mental Demand	2,44	1,5	0,9	25,0	0,299	-0,389
Temporal Demand	4,56	3,8	0,8	23,5	0,430	-0,306
Frustration	4,67	4,3	0,4	20,0	0,810	-0,111

5.1.2.7 LS-TLX – Interviews –First LS vs Not First LS

TLX Metric	LS-CC-f	LS-CC-nf	Mean Difference	W	P-Value	Effect Size
Effort	3,40	4,6	-1,2	13,5	0,240	0,400
Performance	3,00	4,6	-1,6	14,0	0,274	0,378
Mental Demand	3,60	4,7	-1,1	18,0	0,585	0,200
Temporal Demand	2,60	3,1	-0,5	20,5	0,828	0,089
Frustration	1,60	2,4	-0,8	21,0	0,889	0,067

5.1.3 Task Specific Difficulty

5.1.3.1 Written feedback – Starting Participants

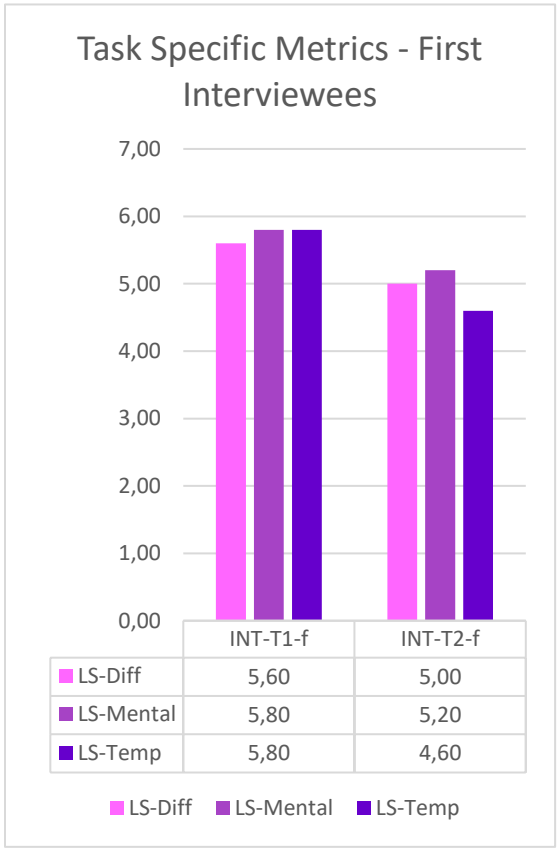
BCM	**CJM**
n/a	lot to cover depending on how detailed you go
none	Difficult to think about expanding any pains and gains really pertaining to the different phases of a project or program. They might be more generic in nature.
.	.
NA	NA
NA	NA

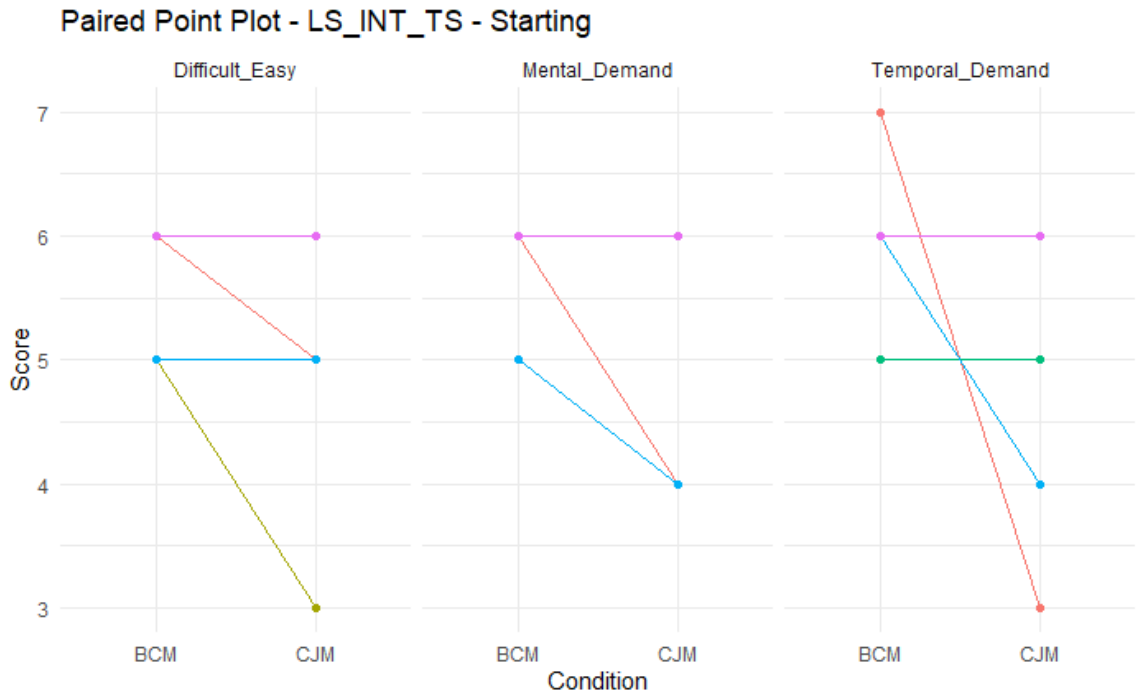
5.1.3.2 Written feedback – Follow-up Participants

BCM	**CJM**
As an analyst you are not so involved in the sell-side of a project	was okay
N/A	N/A
It could have been mor explanatory.	I did not
time	N/a
Bit of duplications but overal OK	Ansers are more generic Accenture WoW instead of LS specific
N/A	The time slot is not enough for this section
N/A	N/A
N/A	N/A
yes it was fast / hurried, but it also has a positive effect on focus, choosing. However, I can imagine for others they need more time to feel comforrable and satisfied with the output	see answer before, and the prep in the journey helped in speeding up

5.1.3.3 LS-TS – Interviews – Starting Participants

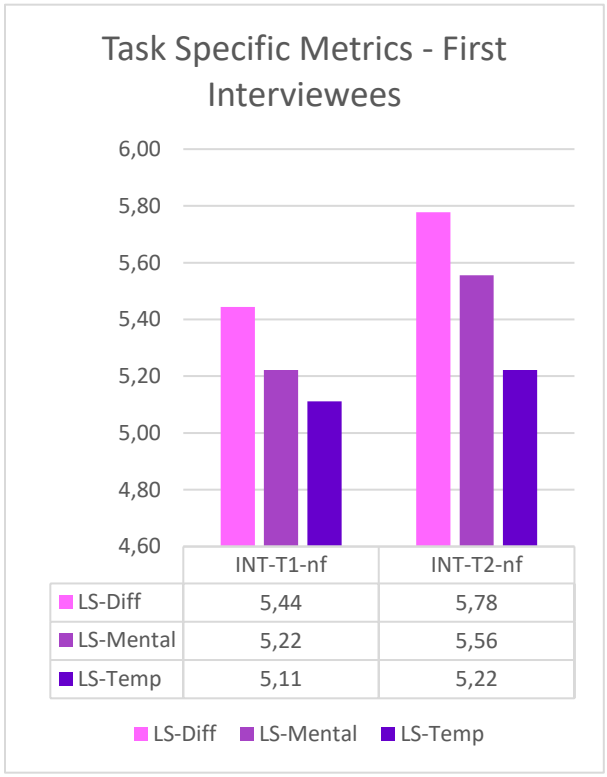
TS Metric	Mean First LS BCM	Mean First LS CJM	Mean Difference	P-Value	Effect Size
Difficulty	5,6	5	0,6	0,488422	-0,28
Mental	5,8	5,2	0,6	0,438578	-0,28
Temporal	5,8	4,6	1,2	0,126379	-0,6

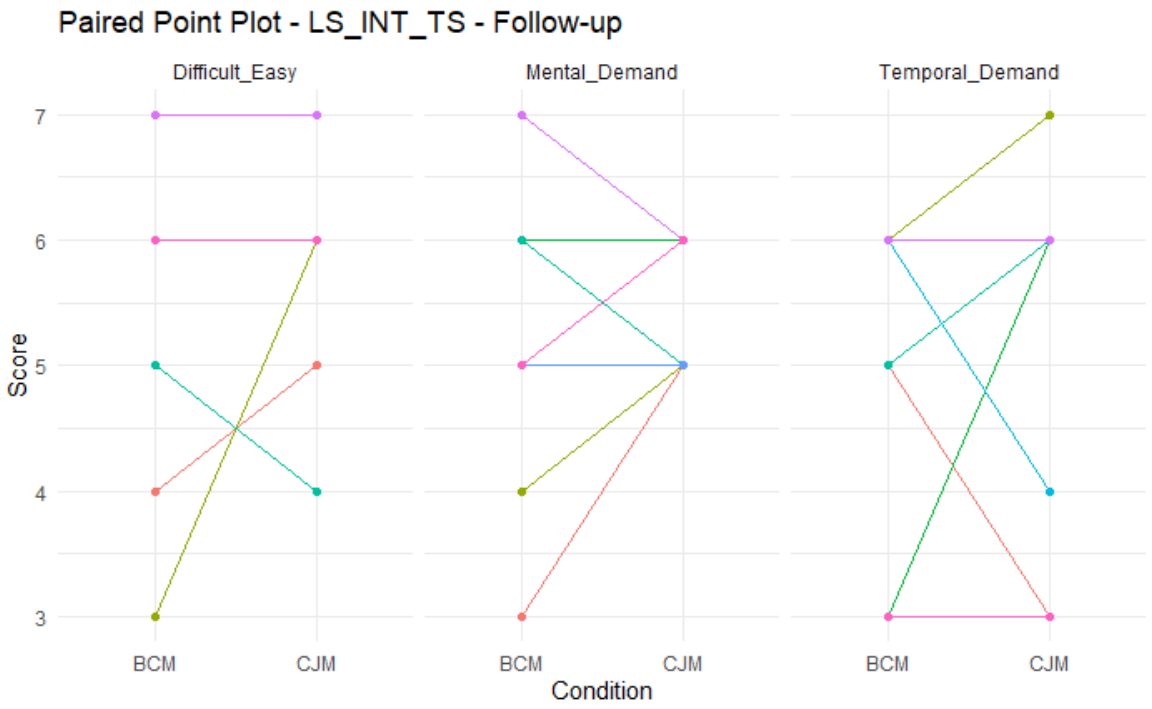




5.1.3.4 LS-TS – Interviews – Follow-up Participants

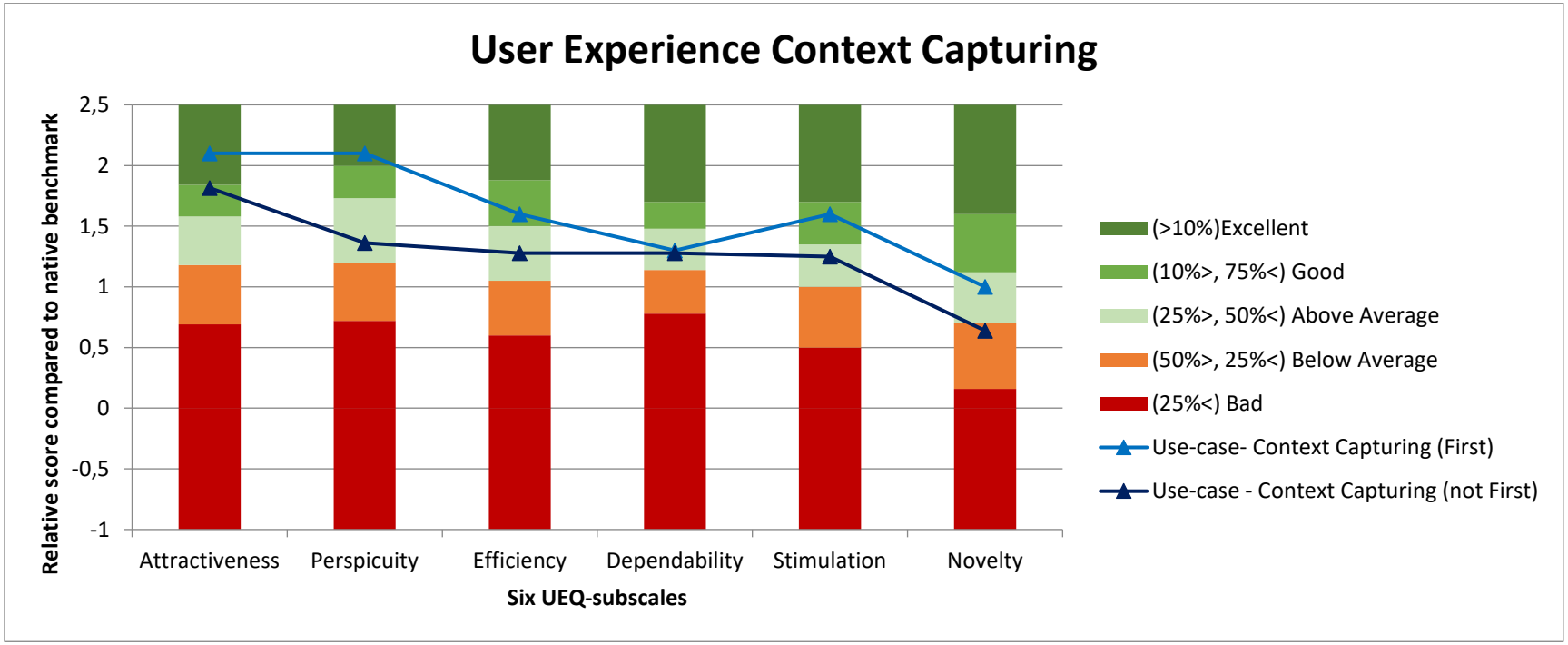
TS Metric	Mean First LS BCM	Mean First LS CJM	Mean Difference	P-Value	Effect Size
Difficulty	5,4	5,8	-0,333	0,651	0,123
Mental	5,2	5,6	-0,333	0,600	0,123
Temporal	5,1	5,2	-0,111	0,660	0,148





5.1.4 User Experience Questionnaire

5.1.4.1 Use-Case – Context Capturing – UEQ - Summary



The Context Capturing session for the first and not first participants scored above average compared to the benchmark. It stands out that the user experience of the first participants was slightly better than that of the not first participants. This is comparable to the slight difference in the TLX metrics. The largest difference seems to be in perspicuity, which, in the sessions with the not-first participants, showed a high variance, indicating mixed opinions from the participants. Zooming in on this metric reveals that the session was quite understandable but that it was a bit confusing and complicated for some participants.

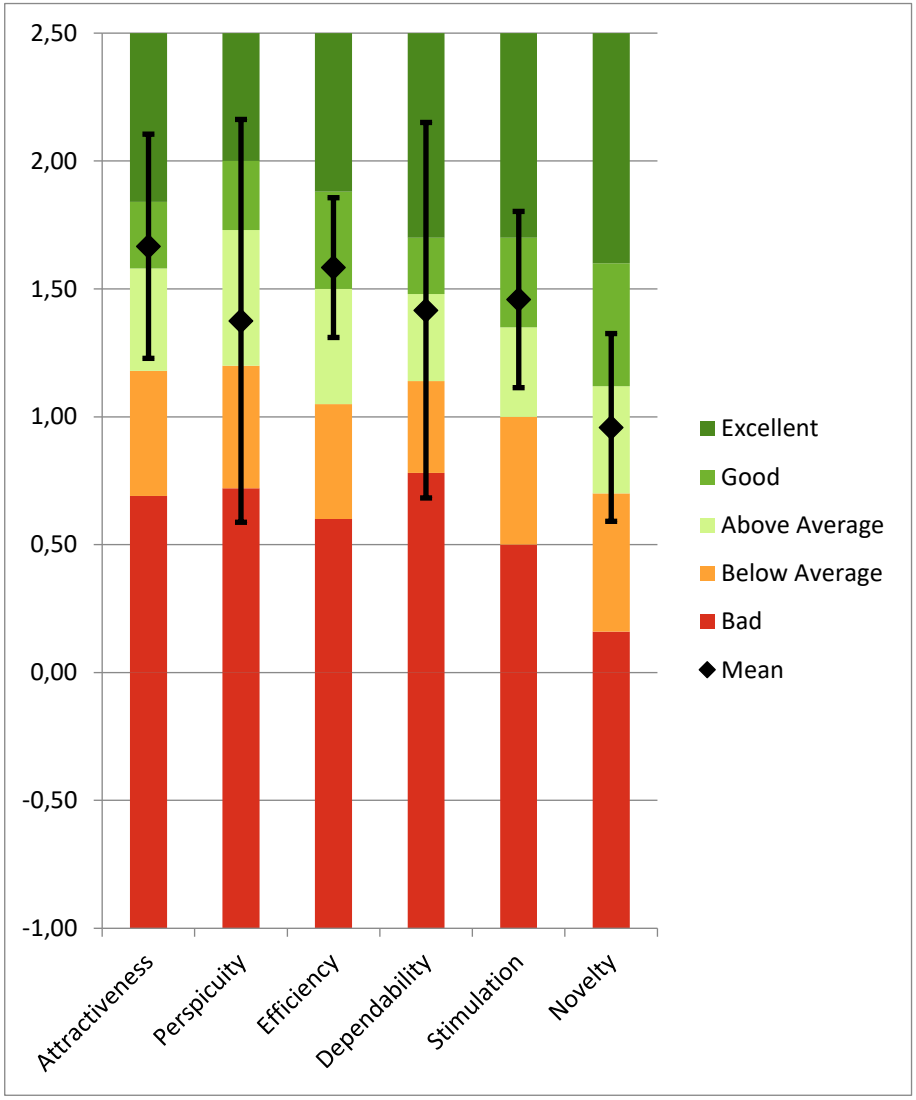
Based on the consistency of the difference in data between the first and non-first participants, we can conclude that the integration of the modified Delphi Method has a slight negative effect on task load or user experience. Overall, however, the Context-Capturing session was well received, showing a reduced task load and proper performance on the UEQ scale.

Most improvements can be made in improving dependability, expectation management, and predictability.

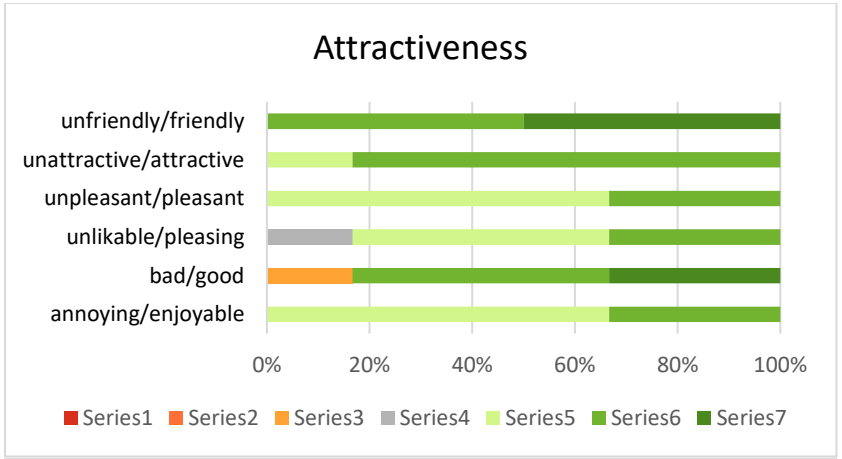
5.1.4.1.1 Starting Participants

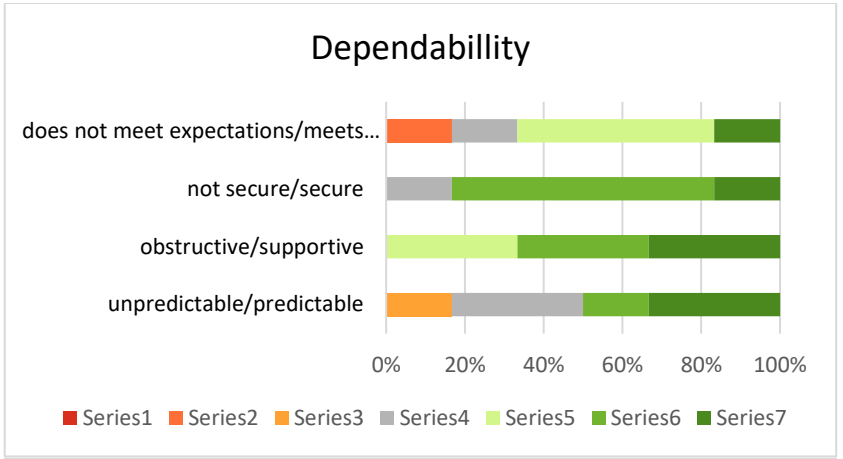
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,67	**Good**	10% of results better, 75% of results worse
Perspicuity	1,38	**Above Average**	25% of results better, 50% of results worse
Efficiency	1,58	**Good**	10% of results better, 75% of results worse
Dependability	1,42	**Above Average**	25% of results better, 50% of results worse
Stimulation	1,46	**Good**	10% of results better, 75% of results worse
Novelty	0,96	**Above Average**	25% of results better, 50% of results worse

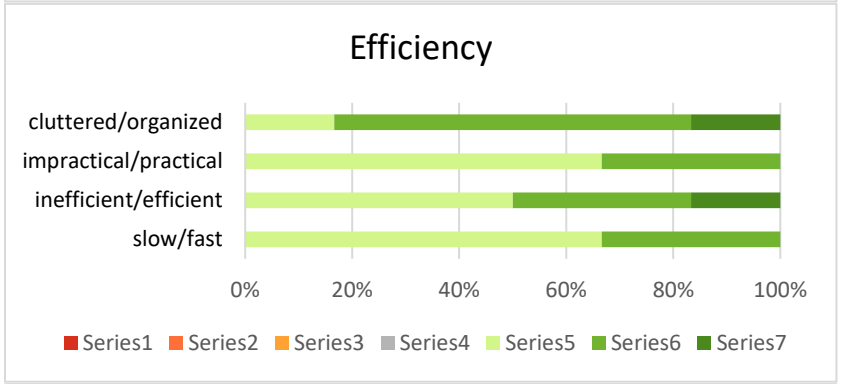
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,667	0,12	0,182	0,81
Perspicuity	1,375	0,08	0,149	-0,29
Efficiency	1,583	0,27	0,272	0,61
Dependability	1,417	0,08	0,143	-1,50
Stimulation	1,458	0,21	0,238	0,48
Novelty	0,958	0,72	0,444	0,94

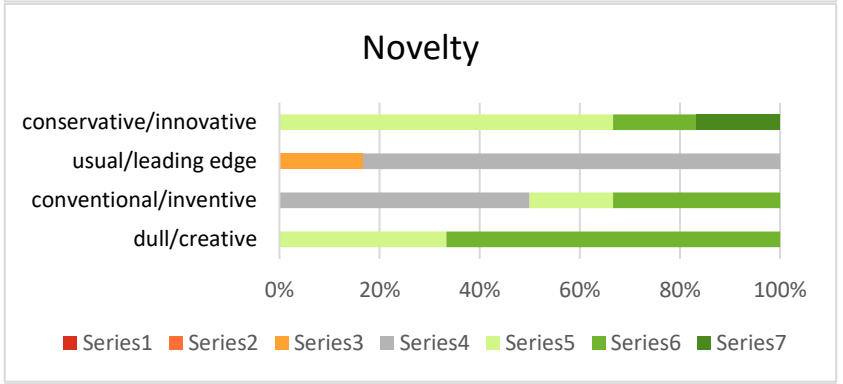


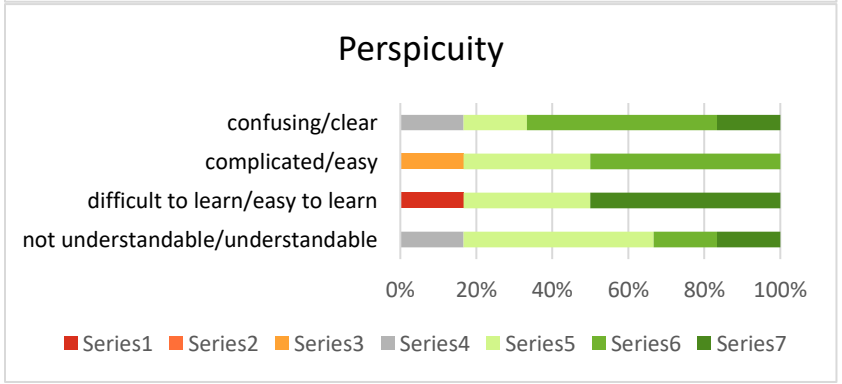
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,3	0,3	0,5	6	annoying	enjoyable	Attractiveness
2	1,3	1,1	1,0	6	not understandable	understandable	Perspicuity
3	1,7	0,3	0,5	6	creative	dull	Novelty
4	1,3	5,5	2,3	6	easy to learn	difficult to learn	Perspicuity
5	1,8	0,6	0,8	6	valuable	inferior	Stimulation
6	0,8	0,6	0,8	6	boring	exciting	Stimulation
7	1,7	0,3	0,5	6	not interesting	interesting	Stimulation
8	1,2	3,0	1,7	6	unpredictable	predictable	Dependability
9	1,3	0,3	0,5	6	fast	slow	Efficiency
10	0,8	1,0	1,0	6	inventive	conventional	Novelty
11	2,0	0,8	0,9	6	obstructive	supportive	Dependability
12	1,8	2,2	1,5	6	good	bad	Attractiveness
13	1,2	1,4	1,2	6	complicated	easy	Perspicuity
14	1,2	0,6	0,8	6	unlikable	pleasing	Attractiveness
15	-0,2	0,2	0,4	6	usual	leading edge	Novelty
16	1,3	0,3	0,5	6	unpleasant	pleasant	Attractiveness
17	1,8	1,0	1,0	6	secure	not secure	Dependability
18	1,5	0,7	0,8	6	motivating	demotivating	Stimulation
19	0,7	2,7	1,6	6	meets expectations	does not meet expectations	Dependability
20	1,7	0,7	0,8	6	inefficient	efficient	Efficiency
21	1,7	1,1	1,0	6	clear	confusing	Perspicuity
22	1,3	0,3	0,5	6	impractical	practical	Efficiency
23	2,0	0,4	0,6	6	organized	cluttered	Efficiency
24	1,8	0,2	0,4	6	attractive	unattractive	Attractiveness
25	2,5	0,3	0,5	6	friendly	unfriendly	Attractiveness
26	1,5	0,7	0,8	6	conservative	innovative	Novelty

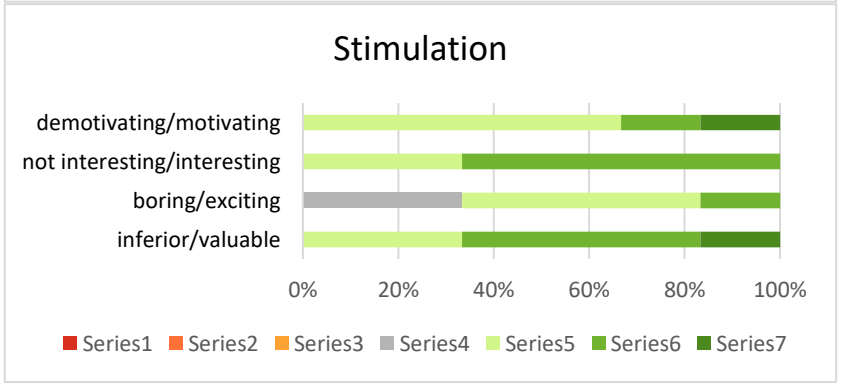

Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7


Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7


Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7


Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7

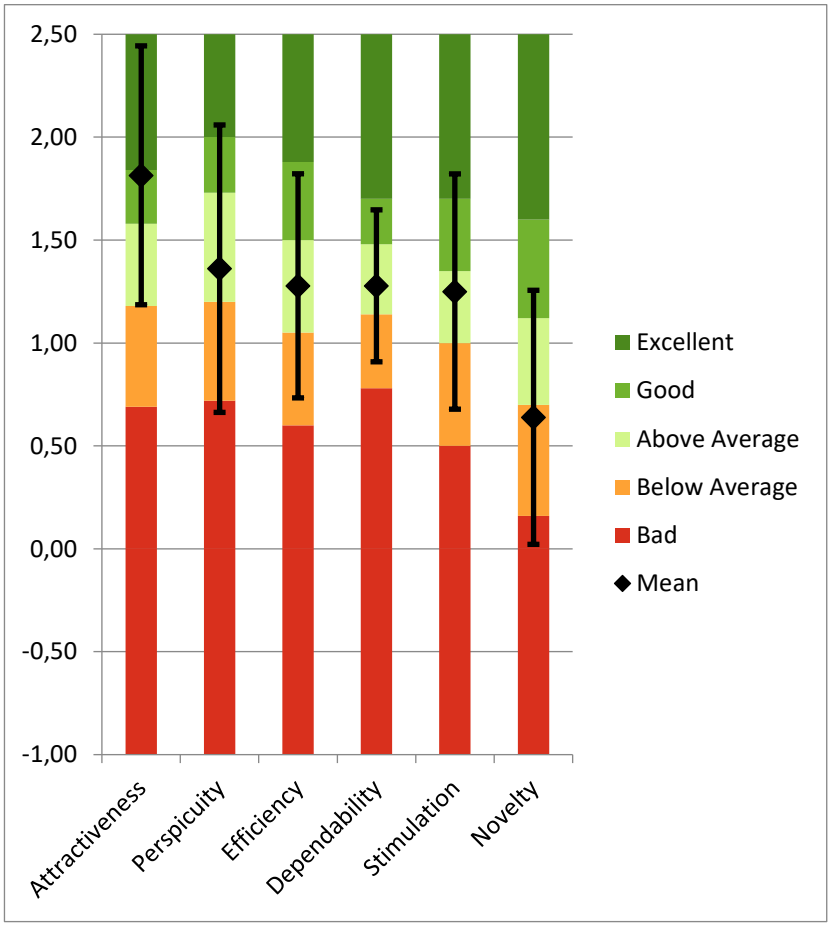

Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7


Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0%
20%
40%
60%
80%
100%
Series1
Series2
Series3
Series4
Series5
Series6
Series7

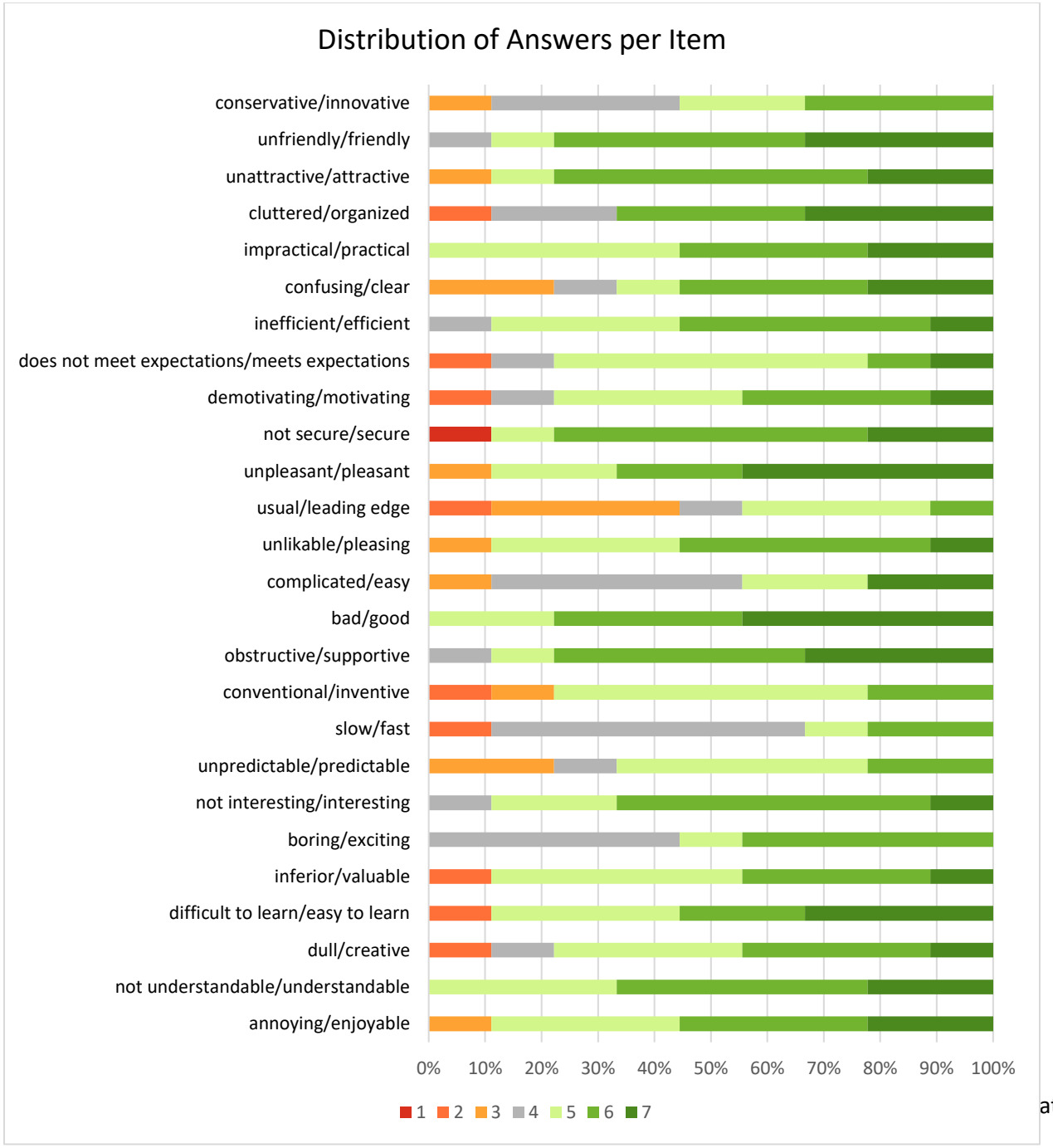
5.1.4.1.2 Follow-up Participants

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,81	**Good**	10% of results better, 75% of results worse
Perspicuity	1,36	**Above Average**	25% of results better, 50% of results worse
Efficiency	1,28	**Above Average**	25% of results better, 50% of results worse
Dependability	1,28	**Above Average**	25% of results better, 50% of results worse
Stimulation	1,25	**Above Average**	25% of results better, 50% of results worse
Novelty	0,64	**Below Average**	50% of results better, 25% of results worse

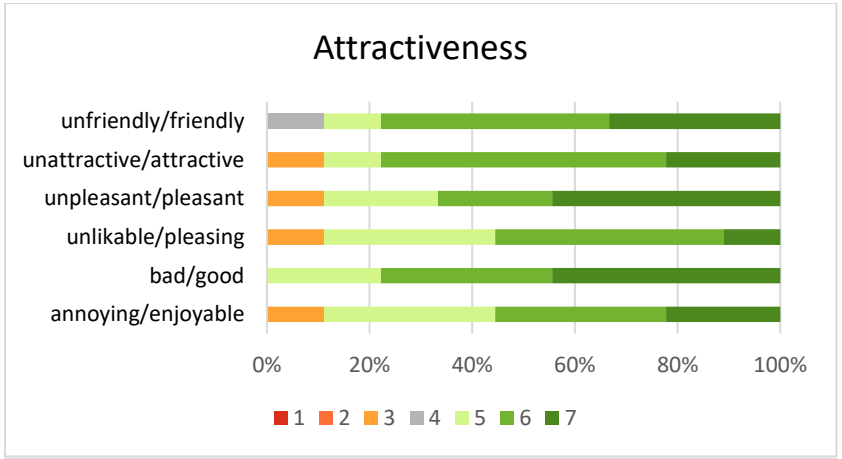
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,815	0,93	0,629	0,92
Perspicuity	1,361	1,14	0,698	0,82
Efficiency	1,278	0,69	0,544	0,67
Dependability	1,278	0,32	0,369	-0,50
Stimulation	1,250	0,77	0,572	0,69
Novelty	0,639	0,89	0,617	0,71

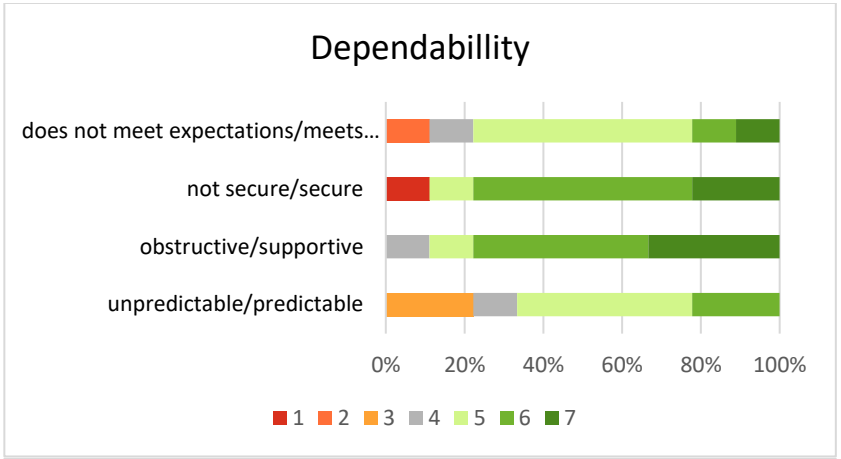


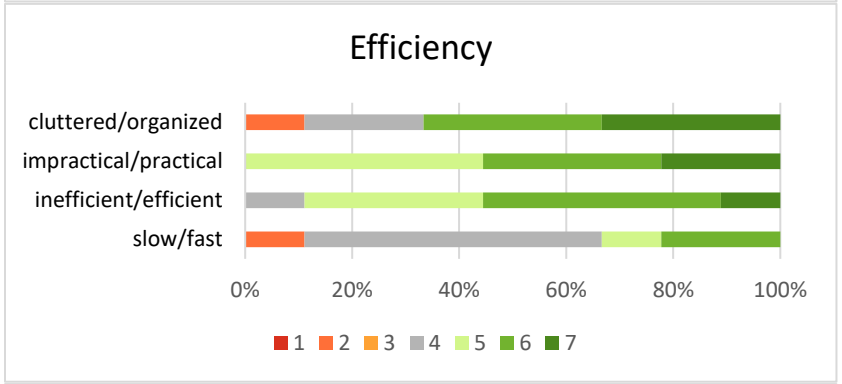
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,6	1,5	1,2	9	annoying	enjoyable	Attractiveness
2	1,9	0,6	0,8	9	not understandable	understandable	Perspicuity
3	1,1	2,1	1,5	9	creative	dull	Novelty
4	1,6	2,5	1,6	9	easy to learn	difficult to learn	Perspicuity
5	1,2	1,9	1,4	9	valuable	inferior	Stimulation
6	1,0	1,0	1,0	9	boring	exciting	Stimulation
7	1,7	0,8	0,9	9	not interesting	interesting	Stimulation
8	0,7	1,3	1,1	9	unpredictable	predictable	Dependability
9	0,3	1,5	1,2	9	fast	slow	Efficiency
10	0,7	1,8	1,3	9	inventive	conventional	Novelty
11	2,0	1,0	1,0	9	obstructive	supportive	Dependability
12	2,2	0,7	0,8	9	good	bad	Attractiveness
13	0,8	1,9	1,4	9	complicated	easy	Perspicuity
14	1,4	1,3	1,1	9	unlikable	pleasing	Attractiveness
15	0,0	1,8	1,3	9	usual	leading edge	Novelty
16	1,9	1,9	1,4	9	unpleasant	pleasant	Attractiveness
17	1,6	3,3	1,8	9	secure	not secure	Dependability
18	1,1	2,1	1,5	9	motivating	demotivating	Stimulation
19	0,9	1,9	1,4	9	meets expectations	does not meet expectations	Dependability
20	1,6	0,8	0,9	9	inefficient	efficient	Efficiency
21	1,2	2,4	1,6	9	clear	confusing	Perspicuity
22	1,8	0,7	0,8	9	impractical	practical	Efficiency
23	1,4	3,0	1,7	9	organized	cluttered	Efficiency
24	1,8	1,4	1,2	9	attractive	unattractive	Attractiveness
25	2,0	1,0	1,0	9	friendly	unfriendly	Attractiveness
26	0,8	1,2	1,1	9	conservative	innovative	Novelty

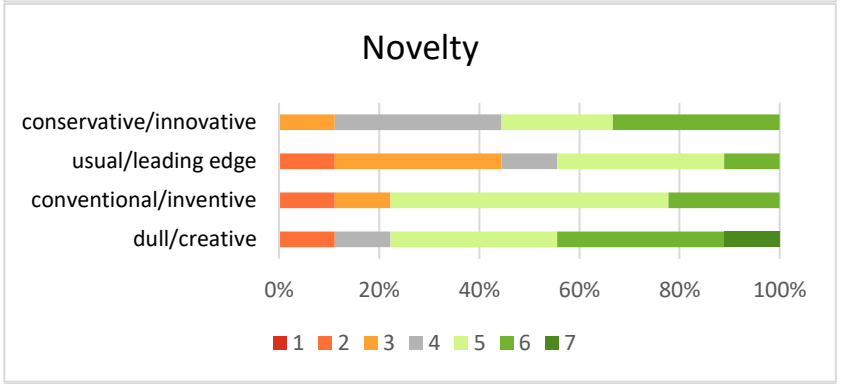


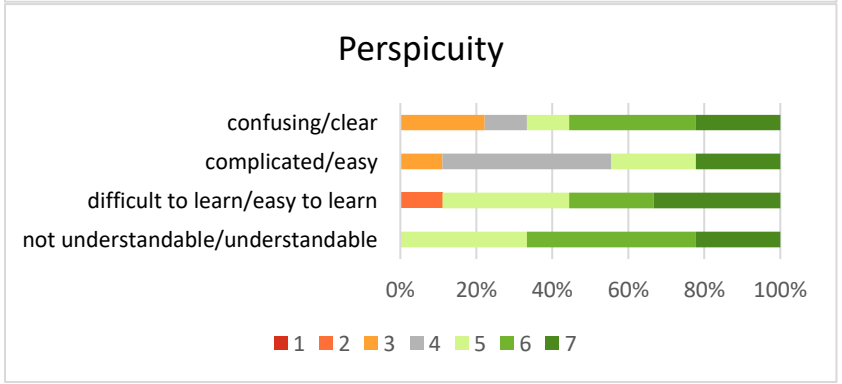
ating Organizational Synergy

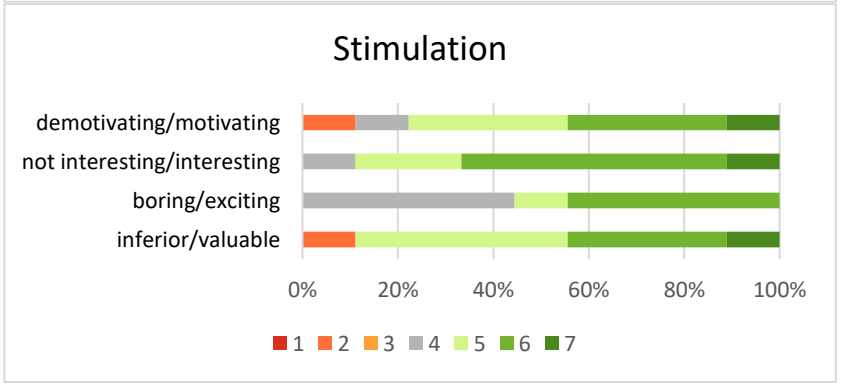

Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7

5.2 Opportunity Identification

5.2.1 End of Survey Comments

Do you have any tips, what could be improved?	Do you have any tops, what went well?	Other feedback?
Less large feedback form	Same	I answered all the same
-	-	-
Could be introduced some groups before workshop. Some people do not know ESP	Good discussion and learning	
Maybe less people or bigger roadmap	The amount of time was enough	Interaction among teams didnt quite happened
To have valuable output on this topic it's crucial to invest time in understanding each other and what the others do. In the prep you might have gained this info but not everybody. It's a lot of info, include more time to understand. Maybe other ways of visualising? It was creative in using different 'tools', keep an eye on the value of shifting constantly. Now we had to understand what is asked from us, the content, work with people we don't know, and taking it away based on words and handwriting. Eithout an explanation of what everybody wrote down. Curious on the outcome. Had to leave on time unfortunately.	Creative way of pulling together the info. Although complex better than a ppt! Liked the writing on the glass. Formats were well prepared. 'Monopoly gameboard like' layout is interesting to use stronger , leverage that as gamification aspect.	
N/A	N/A	
Personally I might have needed some more explanation on parts such as ESP. As well as design layout was sometimes a bit confusing, which line was which phase.	You gathered a good representative group of high level employees. Loved the different exercises within the same context and canvas.	Good luck with graduation! It was an interesting and useful workshop
Explain the blueprint in a bit more detail	Splitting up the workshops in different parts, active workshop	Really nice!
N.a	N.a	
A larger group would have made event more effective. Also, an overview of the various structures in ACN that are represented	Good to promote discussions between participants from different views	A 66-item survey is better done outside the worksop
When there's a lot on the board people think straight away it's a mess and a lot to work through. Also when you have to start adding things it become hard to read very fast, especially if you start adding transparent post-its on top.	Clustering and teaming up to discuss a single opportunity hotspot	Coloring of the swimlanes was not aligned - for example the pre-sales color indicating where the swimlane started had the same color as the swimlane béfore it. So make the arrow/triangle part the same color as the swimlane it points at.

5.2.2 NASA Task Load Index

5.2.2.1 LS-TLX – Opportunity Identification – LS vs Bench

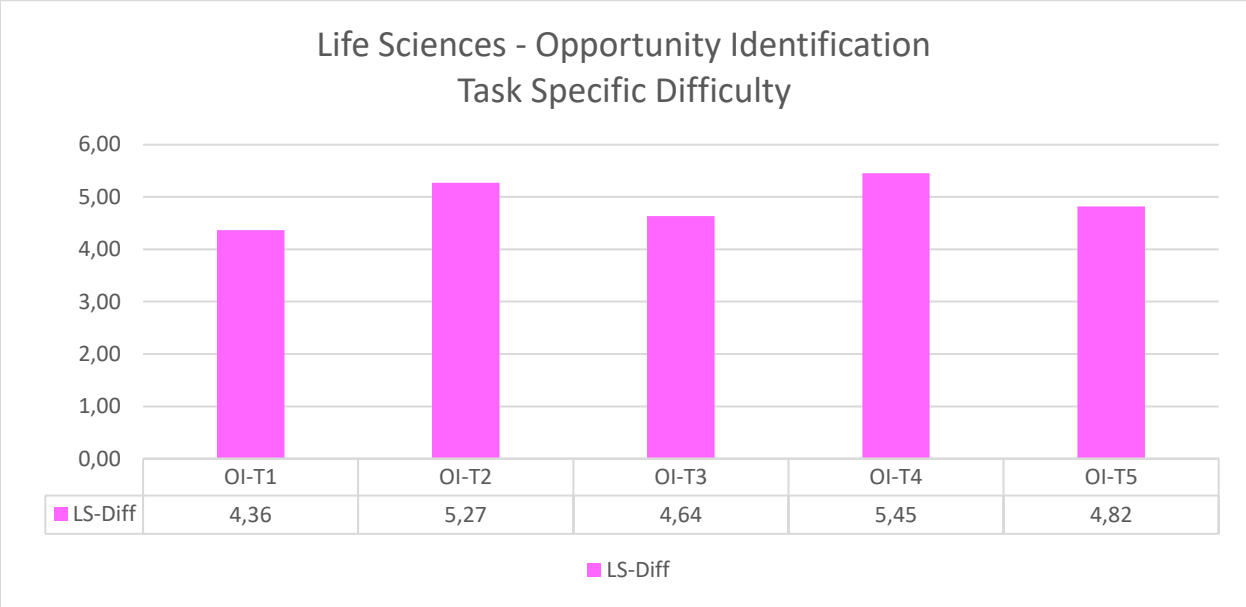
TLX Metric	Mean LS-OI	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,8	6,8	-1,9	4,0	0,016	0,818
Performance	6,1	7,3	-1,2	6,0	0,138	0,523
Mental Demand	5,0	7,0	-2,0	37,0	0,085	0,614
Temporal Demand	3,5	6,3	-2,8	8,5	0,038	0,727
Frustration	4,0	1,0	3,0	10,5	0,053	-0,682

5.2.3 Task Specific Difficulty

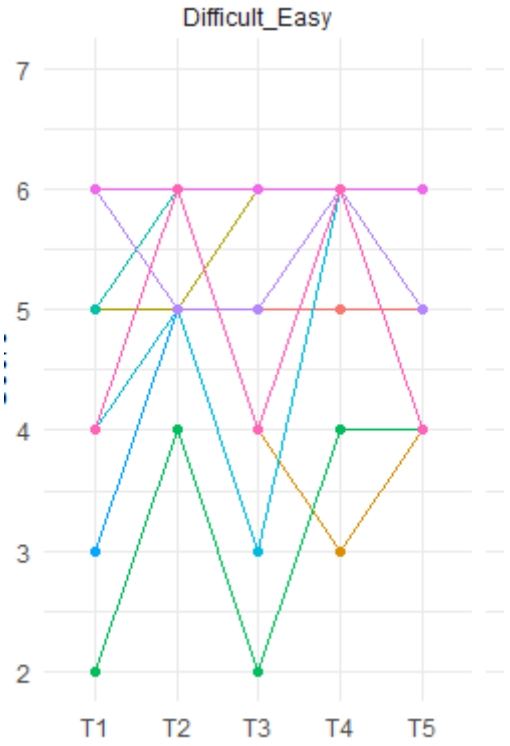
5.2.3.1 Task specific feedback

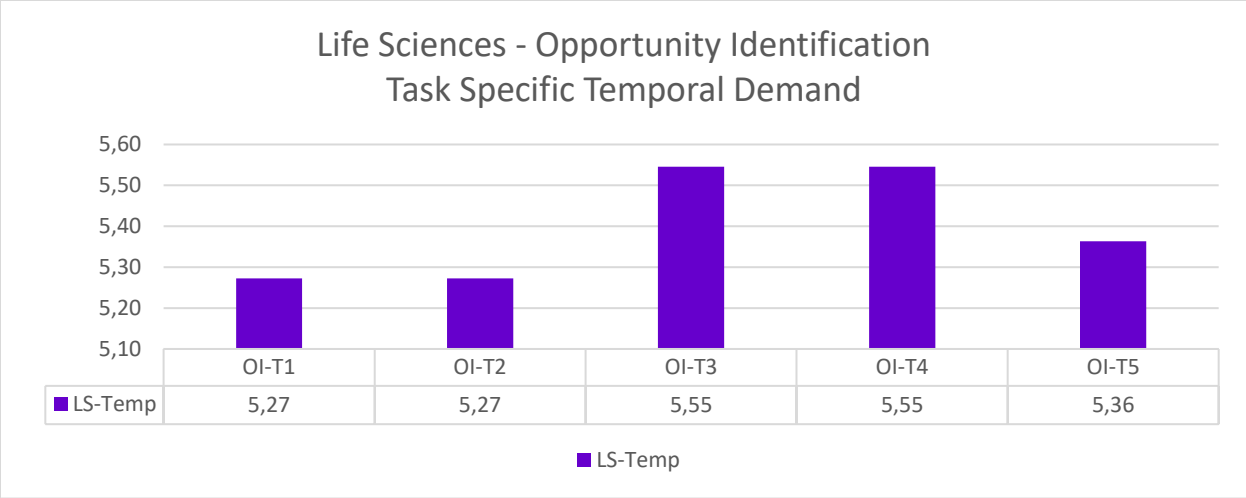
T1	T2	T3	T4	T5
No	No	No	No	No
-	-	-	Readability	Thinking critically
Ok	Not knowing some projects	Ok	Okay	Okay
Too many people around the board and difficult to read	N/A	N/A	N/A	N/A
Complex layout, interpretation of words, unclear what i'm looking for in detail	Felt like the placing was already 'prepared', and more right/wrong exercise. Not sure what I learned from it. Existing collab woth others are unknown for me.	Need to understand what others do to see the match.	-	-
N/A	N/A	N/A	N/A	N/A
-	-	As a recently joined Accenture employee I do not fully understand the organisations processes. Therefore, it is harder to find links to and less common opportunities	-	More in depth vision and creation of research domains is more mentally demanding. As you want a strong and rightfulness fundation
Takes time to understand the blueprint	-	Needs some thinking	Nice!	Nice to think about in a duo

5.2.3.2 Internal Task Comparison

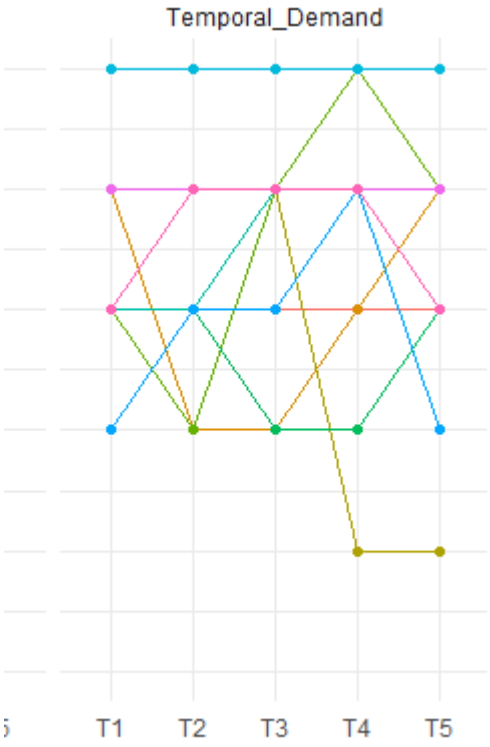


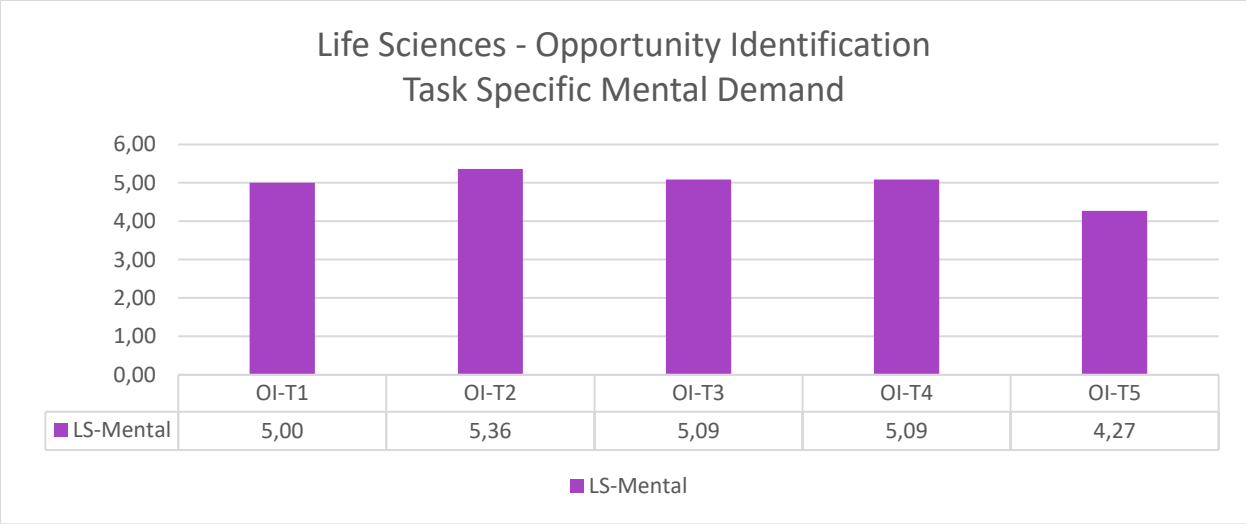
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	T1	T2	4,36	5,27	-0,91	0,08	0,42
Difficulty	T1	T3	4,36	4,64	-0,27	0,61	0,13
Difficulty	T1	T4	4,36	5,45	-1,09	0,03	0,54
Difficulty	T1	T5	4,36	4,82	-0,45	0,49	0,17
Difficulty	T2	T3	5,27	4,64	0,64	0,26	-0,27
Difficulty	T2	T4	5,27	5,45	-0,18	0,26	0,26
Difficulty	T2	T5	5,27	4,82	0,45	0,15	-0,34
Difficulty	T3	T4	4,64	5,45	-0,82	0,07	0,43
Difficulty	T3	T5	4,64	4,82	-0,18	1,00	0,01
Difficulty	T4	T5	5,45	4,82	0,64	0,05	-0,46



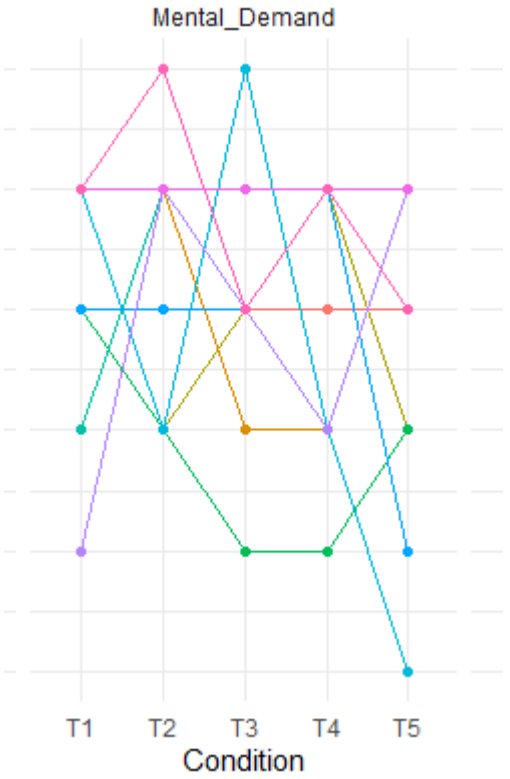


Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Temporal Demand	T1	T2	5,27	5,27	0,00	1,00	0,00
Temporal Demand	T1	T3	5,27	5,55	-0,27	0,42	0,20
Temporal Demand	T1	T4	5,27	5,55	-0,27	0,39	0,21
Temporal Demand	T1	T5	5,27	5,36	-0,09	0,65	0,12
Temporal Demand	T2	T3	5,27	5,55	-0,27	0,42	0,20
Temporal Demand	T2	T4	5,27	5,55	-0,27	0,39	0,21
Temporal Demand	T2	T5	5,27	5,36	-0,09	0,65	0,12
Temporal Demand	T3	T4	5,55	5,55	0,00	0,89	0,04
Temporal Demand	T3	T5	5,55	5,36	0,18	0,75	-0,08
Temporal Demand	T4	T5	5,55	5,36	0,18	0,65	-0,12



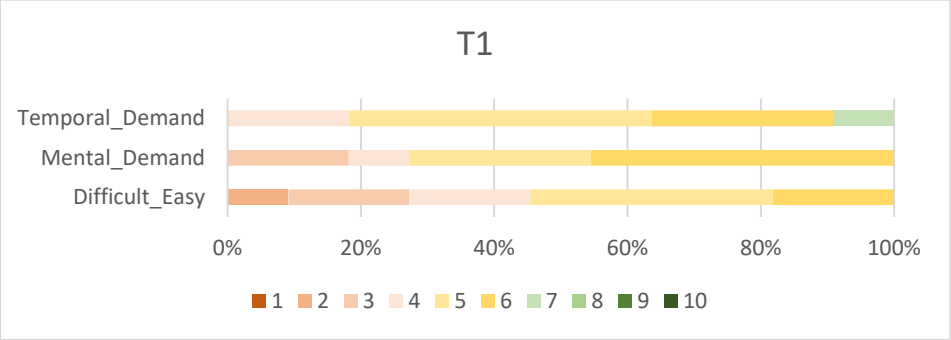


Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Mental Demand	T1	T2	5,00	5,36	-0,36	0,53	0,16
Mental Demand	T1	T3	5,00	5,09	-0,09	0,92	-0,03
Mental Demand	T1	T4	5,00	5,09	-0,09	0,83	0,06
Mental Demand	T1	T5	5,00	4,27	0,73	0,20	-0,32
Mental Demand	T2	T3	5,36	5,09	0,27	0,54	-0,16
Mental Demand	T2	T4	5,36	5,09	0,27	0,64	-0,12
Mental Demand	T2	T5	5,36	4,27	1,09	0,08	-0,44
Mental Demand	T3	T4	5,09	5,09	0,00	0,84	0,06
Mental Demand	T3	T5	5,09	4,27	0,82	0,20	-0,31
Mental Demand	T4	T5	5,09	4,27	0,82	0,16	-0,35

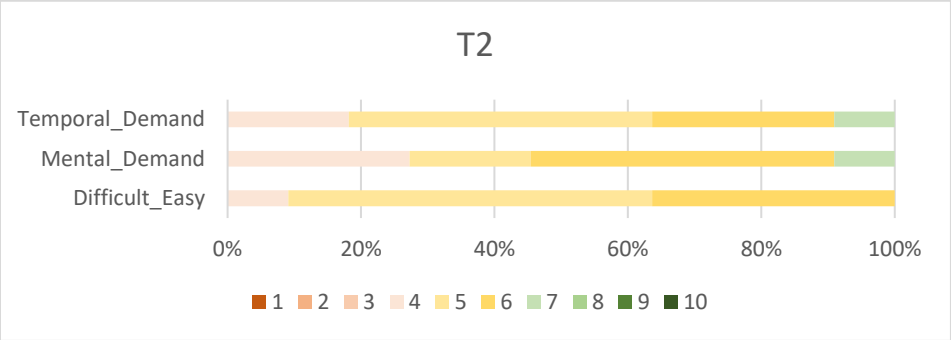


5.2.3.3 Task Specific Data Distribution

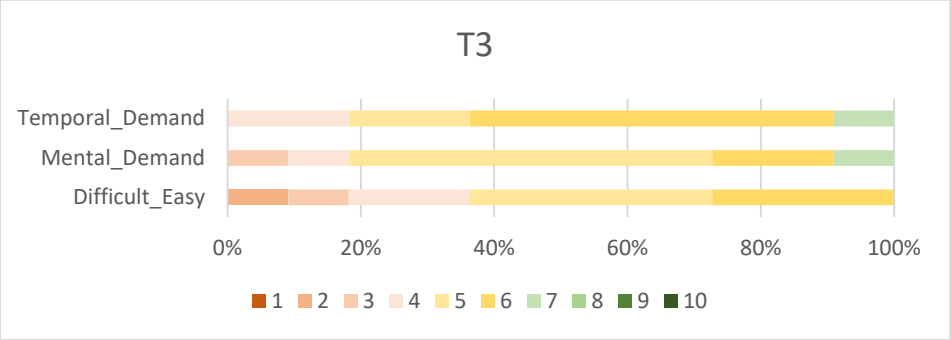
5.2.3.3.1 LS-TS – Opportunity Identification – Task 1



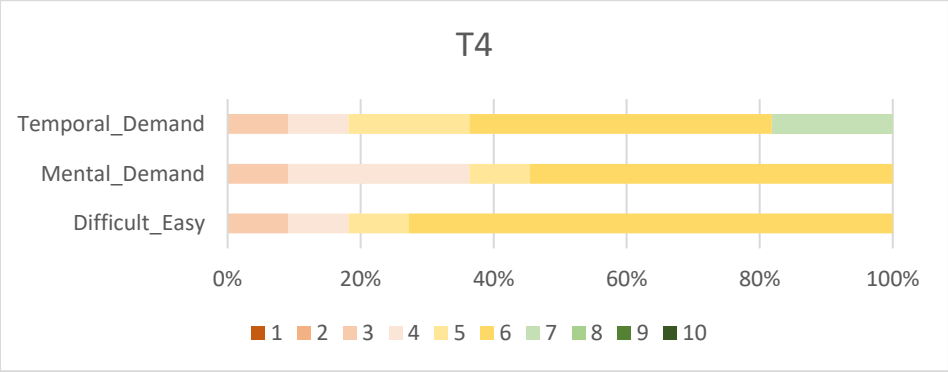
5.2.3.3.2 LS-TS – Opportunity Identification – Task 2



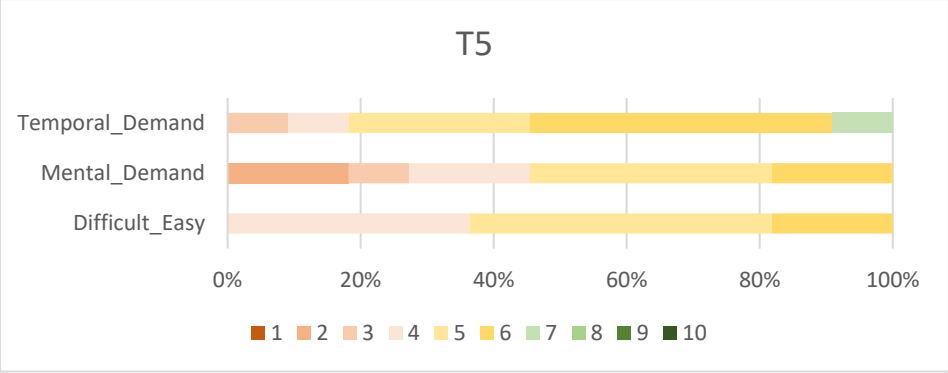
5.2.3.3.3 LS-TS – Opportunity Identification – Task 3



5.2.3.3.4 LS-TS – Opportunity Identification – Task 4



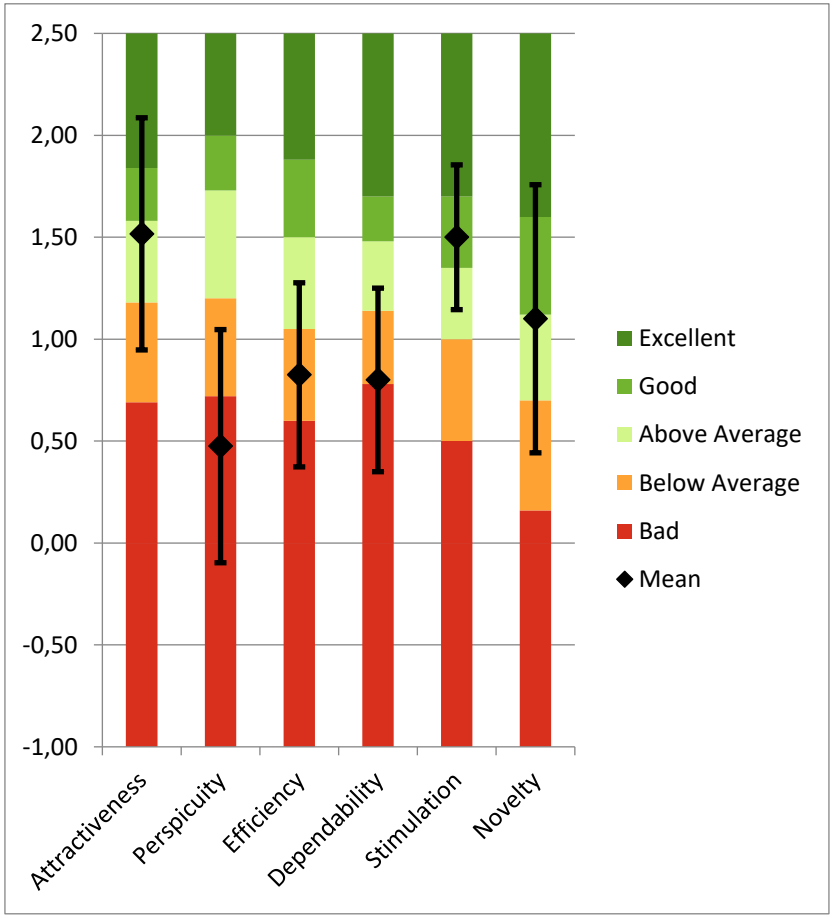
5.2.3.3.5 LS-TS – Opportunity Identification – Task 5



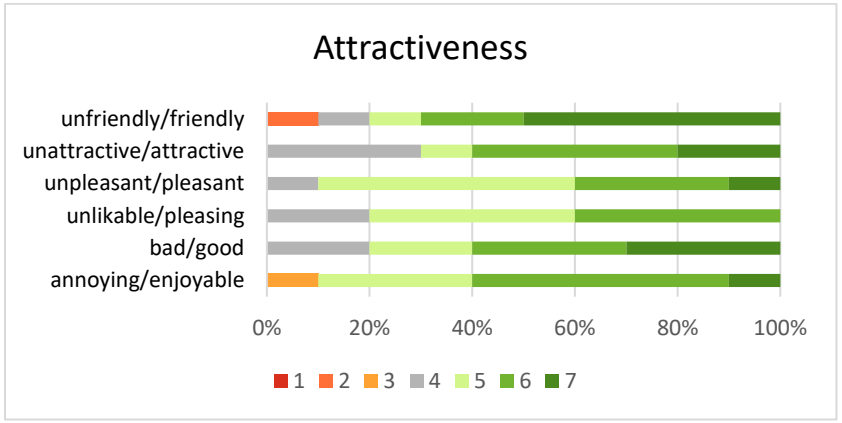
5.2.4 User Experience Questionnaire

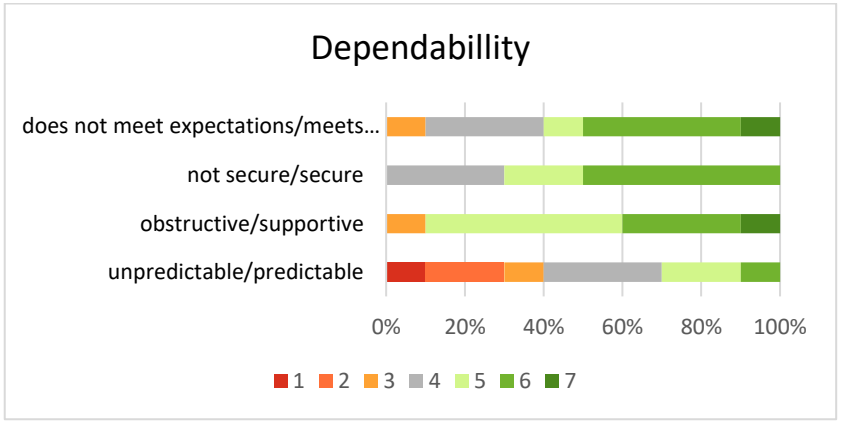
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,52	**Above average**	25% of results better, 50% of results worse
Perspicuity	0,48	**Bad**	In the range of the 25% worst results
Efficiency	0,83	**Below Average**	50% of results better, 25% of results worse
Dependability	0,80	**Below Average**	50% of results better, 25% of results worse
Stimulation	1,50	**Good**	10% of results better, 75% of results worse
Novelty	1,10	**Above Average**	25% of results better, 50% of results worse

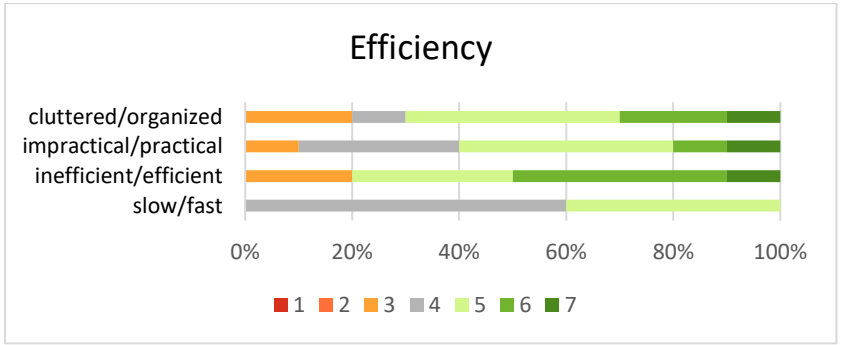
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,517	0,93	0,569	0,93
Perspicuity	0,475	0,94	0,572	0,83
Efficiency	0,825	0,58	0,452	0,43
Dependability	0,800	0,58	0,450	0,40
Stimulation	1,500	0,36	0,355	0,79
Novelty	1,100	1,24	0,658	0,77

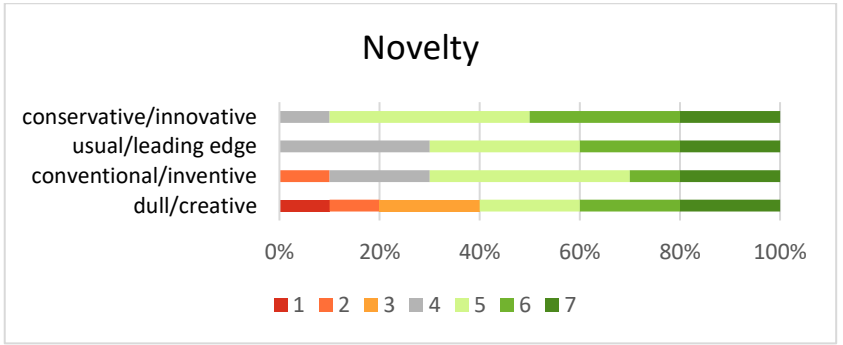


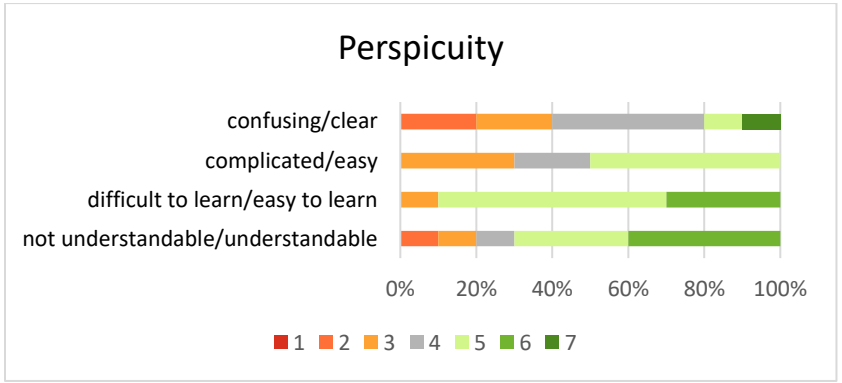
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,5	1,2	1,1	11	annoying	enjoyable	Attractiveness
2	0,8	2,0	1,4	11	not understandable	understandable	Perspicuity
3	0,5	4,5	2,1	11	creative	dull	Novelty
4	1,1	0,8	0,9	11	easy to learn	difficult to learn	Perspicuity
5	1,5	0,5	0,7	11	valuable	inferior	Stimulation
6	1,4	0,5	0,7	11	boring	exciting	Stimulation
7	1,8	0,4	0,6	11	not interesting	interesting	Stimulation
8	-0,4	2,5	1,6	11	unpredictable	predictable	Dependability
9	0,4	0,3	0,5	11	fast	slow	Efficiency
10	1,0	2,2	1,5	11	inventive	conventional	Novelty
11	1,3	1,1	1,1	11	obstructive	supportive	Dependability
12	1,7	1,3	1,2	11	good	bad	Attractiveness
13	0,2	0,8	0,9	11	complicated	easy	Perspicuity
14	1,2	0,6	0,8	11	unlikable	pleasing	Attractiveness
15	1,3	1,3	1,2	11	usual	leading edge	Novelty
16	1,4	0,7	0,8	11	unpleasant	pleasant	Attractiveness
17	1,2	0,8	0,9	11	secure	not secure	Dependability
18	1,3	1,1	1,1	11	motivating	demotivating	Stimulation
19	1,1	1,7	1,3	11	meets expectations	does not meet expectations	Dependability
20	1,2	1,7	1,3	11	inefficient	efficient	Efficiency
21	-0,2	2,2	1,5	11	clear	confusing	Perspicuity
22	0,8	1,3	1,1	11	impractical	practical	Efficiency
23	0,9	1,7	1,3	11	organized	cluttered	Efficiency
24	1,5	1,4	1,2	11	attractive	unattractive	Attractiveness
25	1,8	2,8	1,7	11	friendly	unfriendly	Attractiveness
26	1,6	0,9	1,0	11	conservative	innovative	Novelty

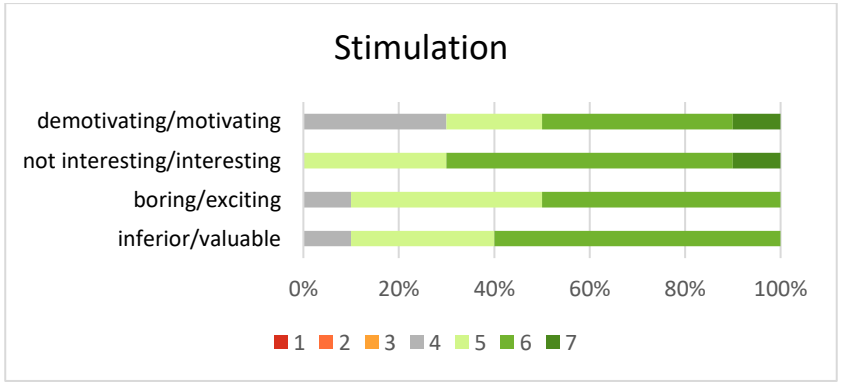

Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7

5.3 Opportunity Selection

5.3.1 End of Survey Comments

Tips	Tops	Other
Could be more people	N/A	
This is a big topic with many aspects, and to do a proper survey would have taken a larger, broader group and more offline time	.	A 63-item survey seems excessive after a 60-minute meeting
balance speed versus content. it now felt more 'a tick in the box' for a graduation project, rather than having time to understand each others work and value and find those content opp's. The questionnaire actually doesn't help in this. I understand you need it for your report, but again, balance..	board was well prepared, with a certain flow in mind.	not sure how and where this happened in the process, but it felt like a 'loop' - the outcome (work together on sales e.g.) is exactly the starting point - how can we integrate/deliver/work more together. I was hoping to get more content driven oppotunities/connections related to client needs.
Slightly less questions, feels like a repetition to confirm :)	-	-

5.3.2 NASA Task Load Index

5.3.2.1 LS-TLX – Opportunity Selection – LS vs Bench

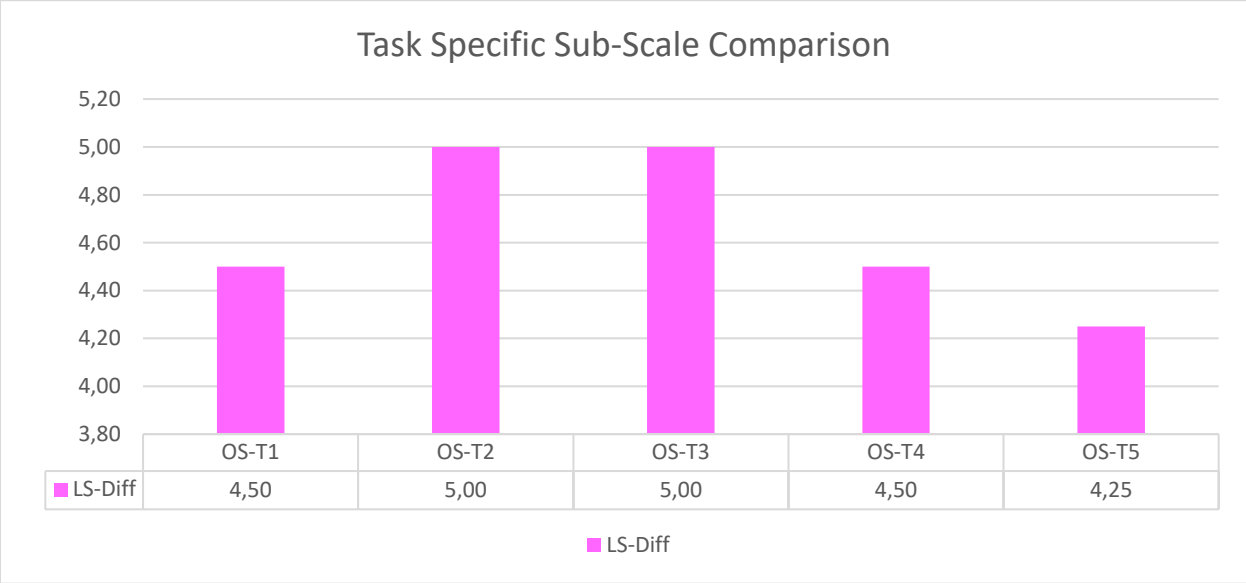
TLX Metric	Mean LS-OS	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,5	6,8	-2,3	0,0	0,080	0,813
Performance	5,0	7,3	-2,3	1,5	0,137	0,688
Mental Demand	4,5	7,0	-2,5	2,5	0,028	1,000
Temporal Demand	5,0	6,3	-1,3	12,5	0,557	0,313
Frustration	4,0	1,3	2,7	5,5	0,234	-0,563

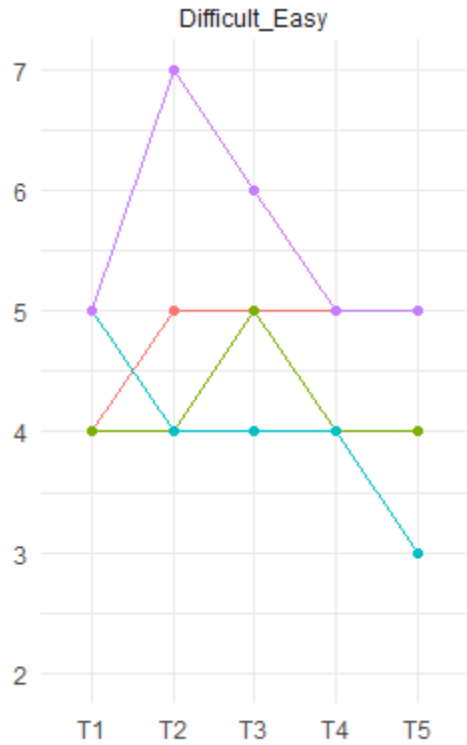
5.3.3 Task Specific Difficulty

5.3.3.1 Task specific feedback

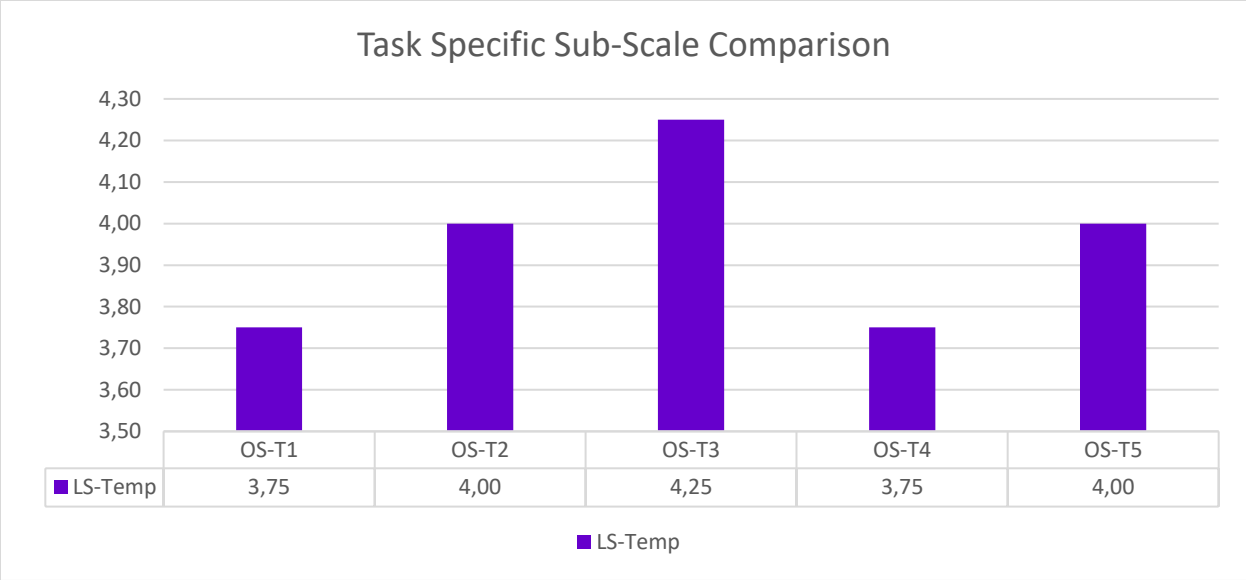
T1	**T2**	**T3**	**T4**	**T5**
need to focus mentally	N/A	N/A	N/A	N/A
I missed this part of the meeting	I stepped in late into the meeting	There was little time to clarify the definitions of opportunities and for a proper comparison	The tight time boxing left insufficient opportunities to discuss with participants	The tight time boxing left insufficient opportunities to discuss with participants
it's just titles - making assumptions on what is meant, discussion time limited	felt double with the next step - clustering and dependencies would be maybe 'more logic'. And it all went 'back' to the sequence of the blueprint. At the end not enough time to do a cross check with everyone - are we all 'happy', really? do we understand why each hexagon is on each place?	again, this takes time for a group to align and understand - why did you place this in this quadrant? , are we sure then that this is 'it'? what are we missing? During the session colleagues needed a definition of the axes. Most are used to setting the priorities and timeline based on other dimensions, such as value/effort. Urgent / important... Then they can be plotted on a time line and	felt a bit double with the previous excercise	was at the end of the meeting - and since we had 'opportunities' that were related to internal, it was somewhat confusing on what was asked for. I see names added, was this done after the workshop? Would have been nice to do this with the content offering, e.g. GenAI or datadriven product development.

5.3.3.2 Internal Task Comparison

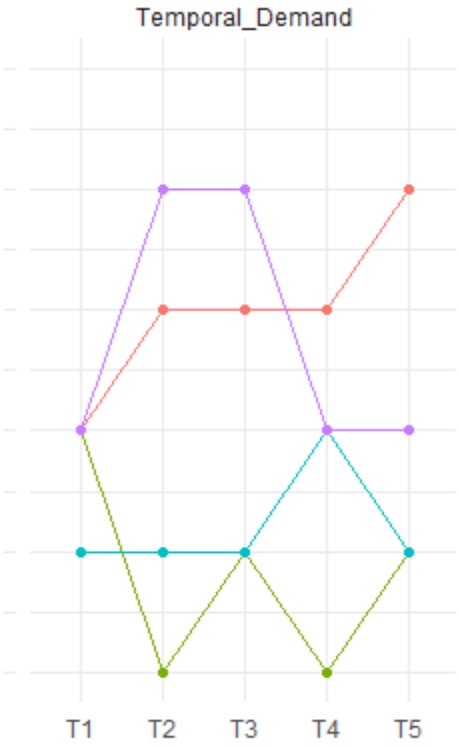


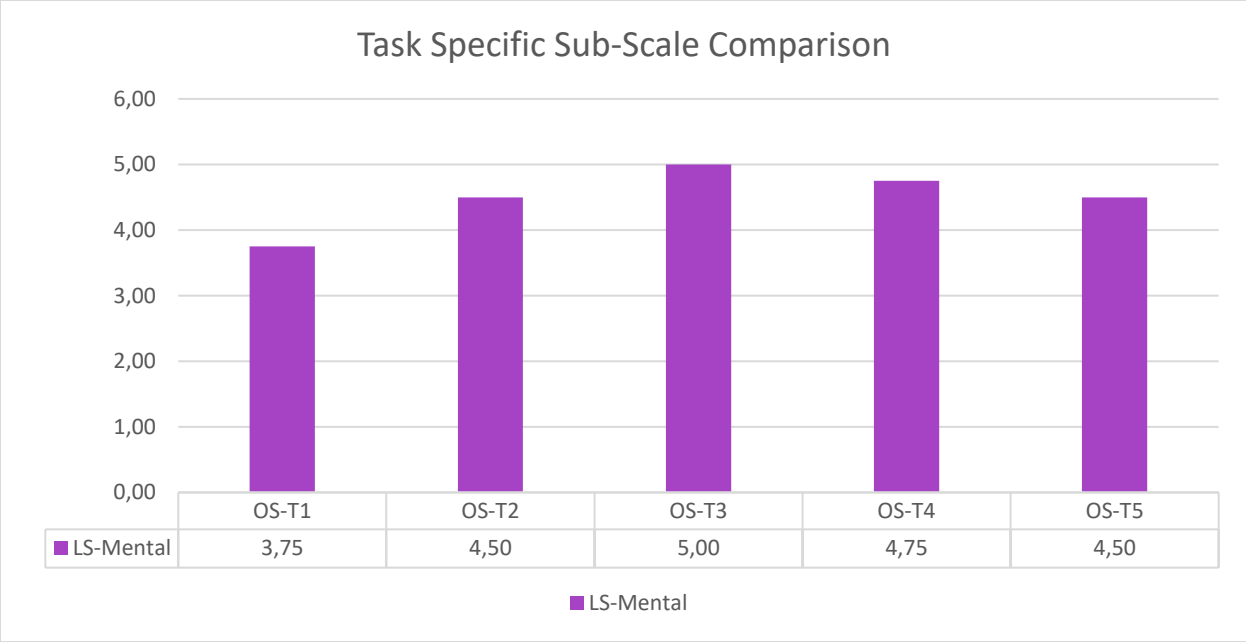


Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Difficulty	T1	T2	4,50	5,00	-0,50	0,87	0,13
Difficulty	T1	T3	4,50	5,00	-0,50	0,43	0,38
Difficulty	T1	T4	4,50	4,50	0,00	1,00	0,00
Difficulty	T1	T5	4,50	4,25	0,25	0,87	-0,13
Difficulty	T2	T3	5,00	5,00	0,00	0,88	0,13
Difficulty	T2	T4	5,00	4,50	0,50	0,87	-0,13
Difficulty	T2	T5	5,00	4,25	0,75	0,65	-0,25
Difficulty	T3	T4	5,00	4,50	0,50	0,43	-0,38
Difficulty	T3	T5	5,00	4,25	0,75	0,35	-0,44
Difficulty	T4	T5	4,50	4,25	0,25	0,87	-0,13

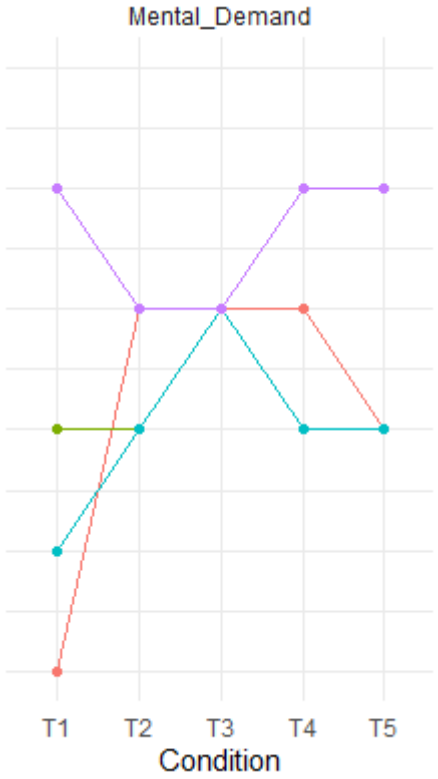


Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Temporal Demand	T1	T2	3,75	4,00	-0,25	1,00	0,06
Temporal Demand	T1	T3	3,75	4,25	-0,50	0,88	0,13
Temporal Demand	T1	T4	3,75	3,75	0,00	0,87	0,13
Temporal Demand	T1	T5	3,75	4,00	-0,25	1,00	-0,06
Temporal Demand	T2	T3	4,00	4,25	-0,25	0,88	0,13
Temporal Demand	T2	T4	4,00	3,75	0,25	0,88	-0,13
Temporal Demand	T2	T5	4,00	4,00	0,00	1,00	0,06
Temporal Demand	T3	T4	4,25	3,75	0,50	0,77	-0,19
Temporal Demand	T3	T5	4,25	4,00	0,25	1,00	-0,06
Temporal Demand	T4	T5	3,75	4,00	-0,25	1,00	0,00



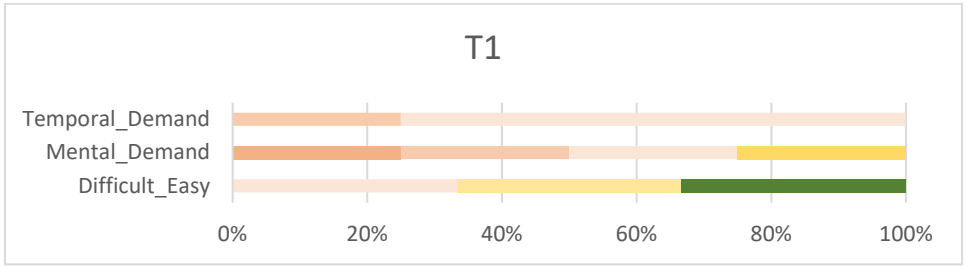


Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect Size
Mental Demand	T1	T2	3,75	4,50	-0,75	0,46	0,38
Mental Demand	T1	T3	3,75	5,00	-1,25	0,28	0,50
Mental Demand	T1	T4	3,75	4,75	-1,00	0,37	0,44
Mental Demand	T1	T5	3,75	4,50	-0,75	0,44	0,38
Mental Demand	T2	T3	4,50	5,00	-0,50	0,18	0,50
Mental Demand	T2	T4	4,50	4,75	-0,25	0,87	0,13
Mental Demand	T2	T5	4,50	4,50	0,00	0,87	-0,13
Mental Demand	T3	T4	5,00	4,75	0,25	0,62	-0,25
Mental Demand	T3	T5	5,00	4,50	0,50	0,27	-0,50
Mental Demand	T4	T5	4,75	4,50	0,25	0,74	-0,19

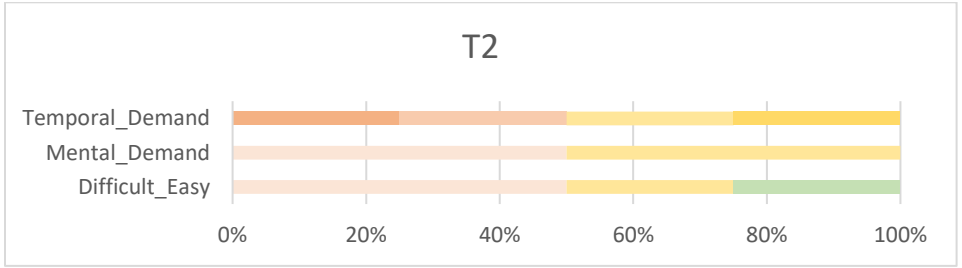


5.3.3.3 Task Specific Data Distribution

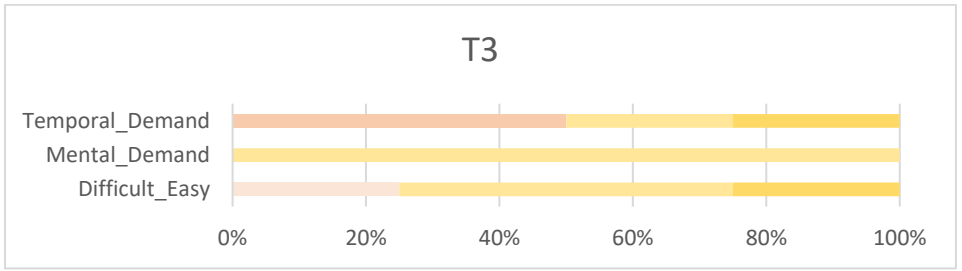
5.3.3.3.1 LS-TS – Opportunity Selection – Task 1



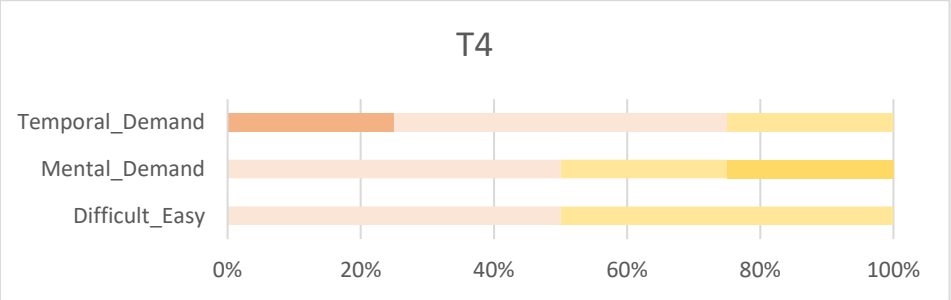
5.3.3.3.2 LS-TS – Opportunity Selection – Task 2



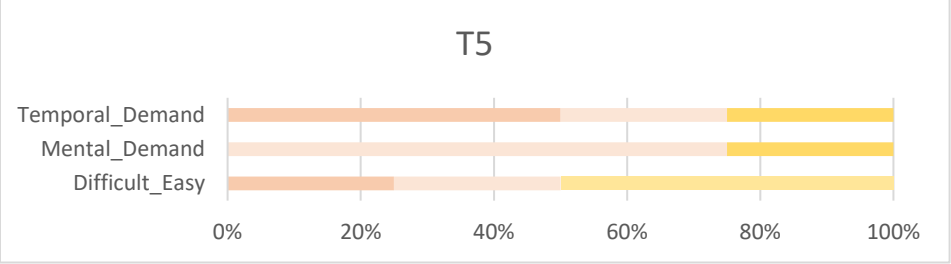
5.3.3.3.3 LS-TS – Opportunity Selection – Task 3



5.3.3.3.4 LS-TS – Opportunity Selection – Task 4



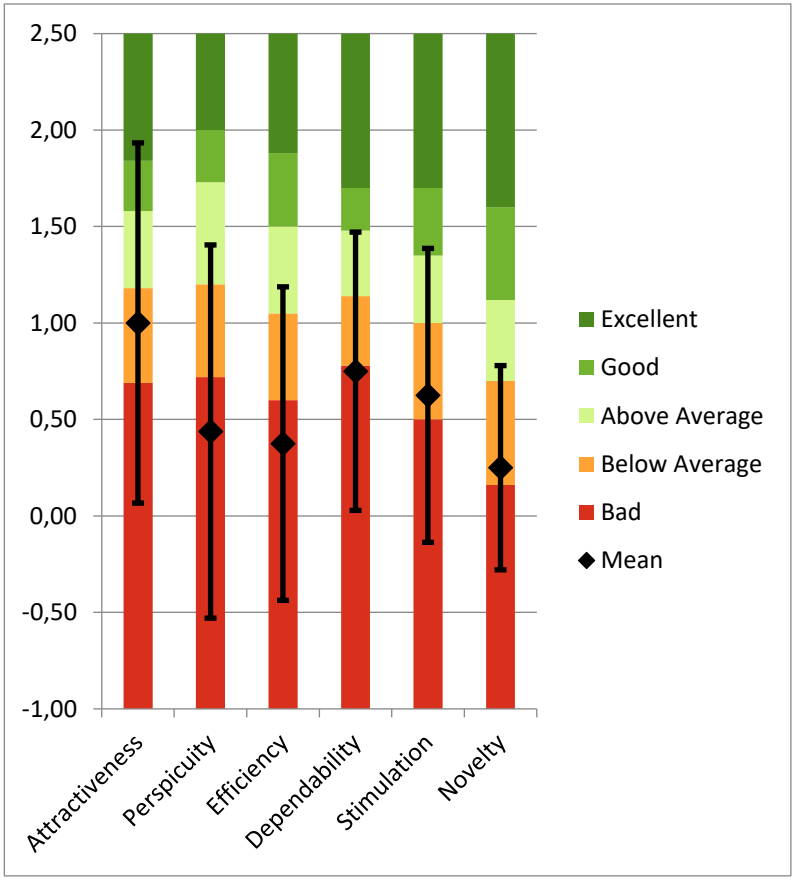
5.3.3.3.5 LS-TS – Opportunity Selection – Task 5



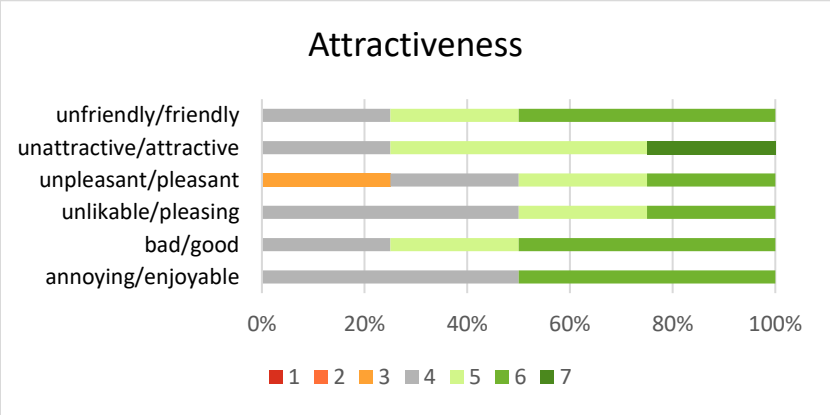
5.3.4 User Experience Questionnaire

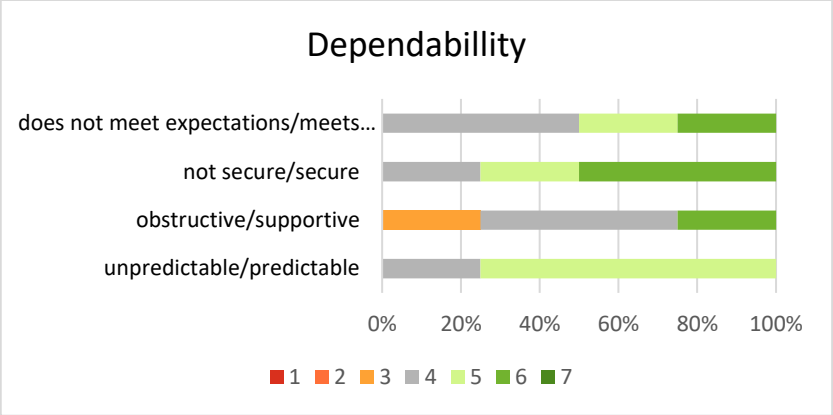
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,00	**Below average**	50% of results better, 25% of results worse
Perspicuity	0,44	**Bad**	In the range of the 25% worst results
Efficiency	0,38	**Bad**	In the range of the 25% worst results
Dependability	0,75	**Bad**	In the range of the 25% worst results
Stimulation	0,63	**Below Average**	50% of results better, 25% of results worse
Novelty	0,25	**Below Average**	50% of results better, 25% of results worse

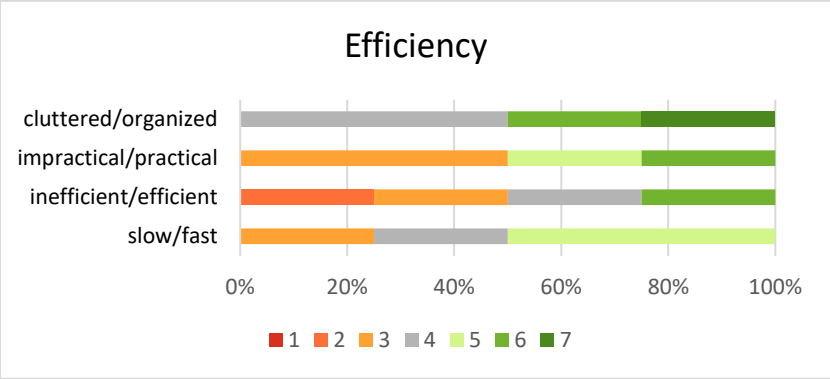
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,000	0,91	0,934	0,93
Perspicuity	0,438	0,97	0,967	0,92
Efficiency	0,375	0,69	0,813	-0,23
Dependability	0,750	0,54	0,721	0,82
Stimulation	0,625	0,60	0,762	0,95
Novelty	0,250	0,29	0,529	-0,79

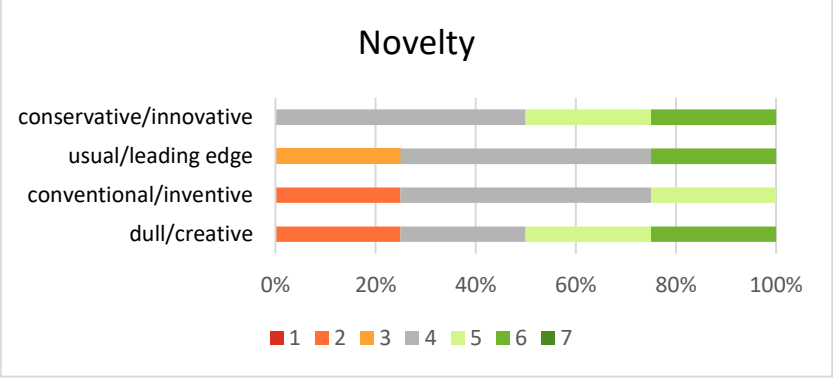


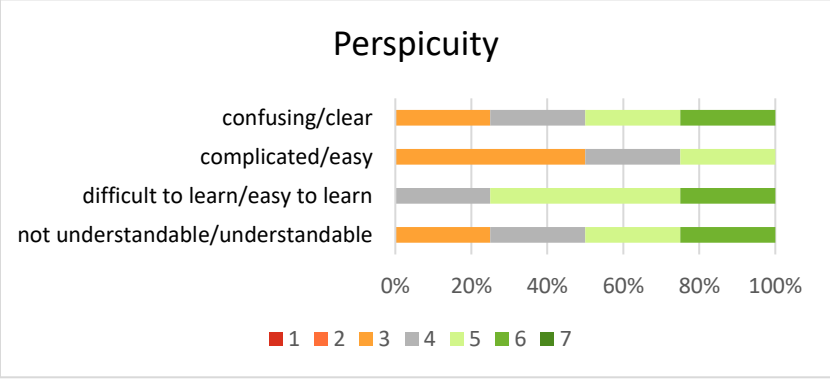
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,0	1,3	1,2	4	annoying	enjoyable	Attractiveness
2	0,5	1,7	1,3	4	not understandable	understandable	Perspicuity
3	0,3	2,9	1,7	4	creative	dull	Novelty
4	1,0	0,7	0,8	4	easy to learn	difficult to learn	Perspicuity
5	0,8	0,9	1,0	4	valuable	inferior	Stimulation
6	0,3	0,3	0,5	4	boring	exciting	Stimulation
7	0,8	0,9	1,0	4	not interesting	interesting	Stimulation
8	0,8	0,3	0,5	4	unpredictable	predictable	Dependability
9	0,3	0,9	1,0	4	fast	slow	Efficiency
10	-0,3	1,6	1,3	4	inventive	conventional	Novelty
11	0,3	1,6	1,3	4	obstructive	supportive	Dependability
12	1,3	0,9	1,0	4	good	bad	Attractiveness
13	-0,3	0,9	1,0	4	complicated	easy	Perspicuity
14	0,8	0,9	1,0	4	unlikable	pleasing	Attractiveness
15	0,3	1,6	1,3	4	usual	leading edge	Novelty
16	0,5	1,7	1,3	4	unpleasant	pleasant	Attractiveness
17	1,3	0,9	1,0	4	secure	not secure	Dependability
18	0,8	0,9	1,0	4	motivating	demotivating	Stimulation
19	0,8	0,9	1,0	4	meets expectations	does not meet expectations	Dependability
20	-0,3	2,9	1,7	4	inefficient	efficient	Efficiency
21	0,5	1,7	1,3	4	clear	confusing	Perspicuity
22	0,3	2,3	1,5	4	impractical	practical	Efficiency
23	1,3	2,3	1,5	4	organized	cluttered	Efficiency
24	1,3	1,6	1,3	4	attractive	unattractive	Attractiveness
25	1,3	0,9	1,0	4	friendly	unfriendly	Attractiveness
26	0,8	0,9	1,0	4	conservative	innovative	Novelty

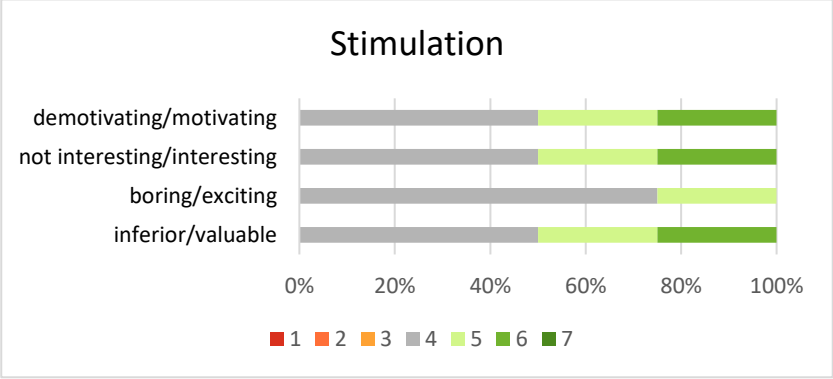

Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7

5.4 Opportunity Modelling

5.4.1 End of Survey Comments

Do you have any tips?	Do you have any tops?	Other feedback?
If possible do a session like this in person; provide some prep for the participants to make the process more smooth	Innovative, creative and fun way of mapping a problem	Thanks for the invitation!
Considering the short time available, this task could benefit from a more prepared set of statements with a request for comments/deltas	.	Focusing only on prio topics has the risk of disregarding others
maybe like it was said to share some topics to think about before the session	N/A	
try not to overcomplicate	very well structured	

5.4.2 NASA Task Load Index

5.4.2.1 LS-TLX – Opportunity Modelling – LS vs Bench

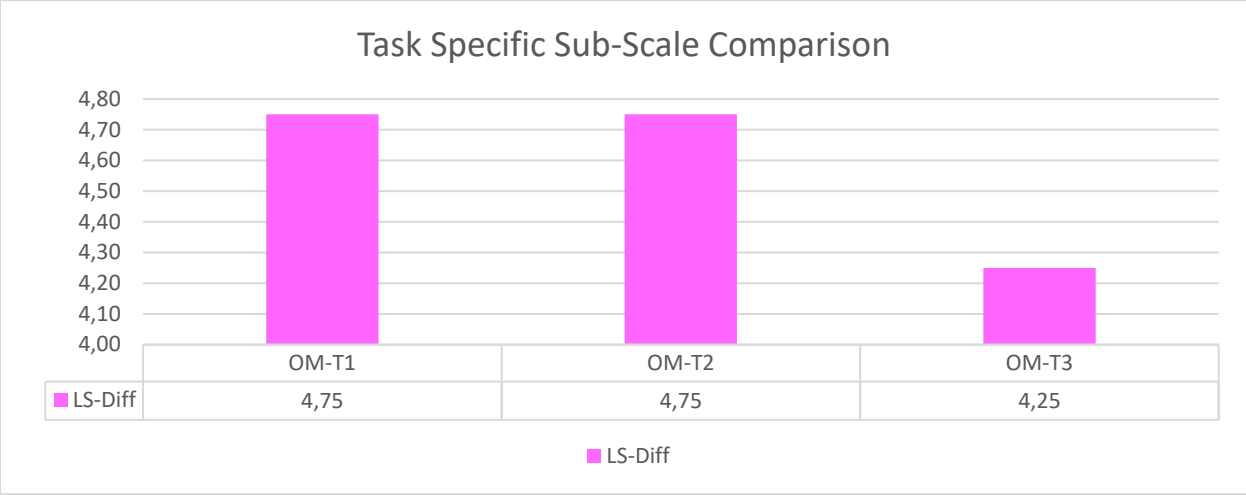
TLX Metric	Mean LS-OM	Mean Benchmark	Mean Difference	W	P-Value	Effect Size
Effort	4,0	6,8	-2,8	-2.75	0,026	1,000
Performance	5,0	7,3	-2,3	-2.25	0,078	0,813
Mental Demand	6,0	7,0	-1,0	-1.00	0,301	0,500
Temporal Demand	5,0	6,3	-1,3	-1.25	0,372	0,438
Frustration	2,8	1,3	1,5	1.75	0,457	-0,375

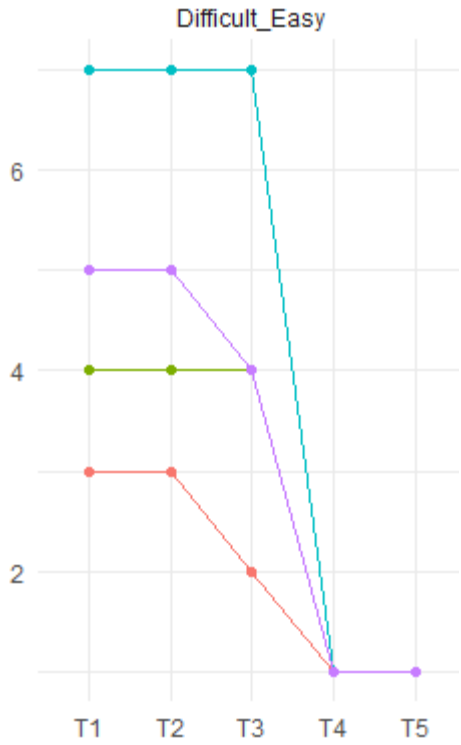
5.4.3 Task Specific Difficulty

5.4.3.1 Task specific feedback

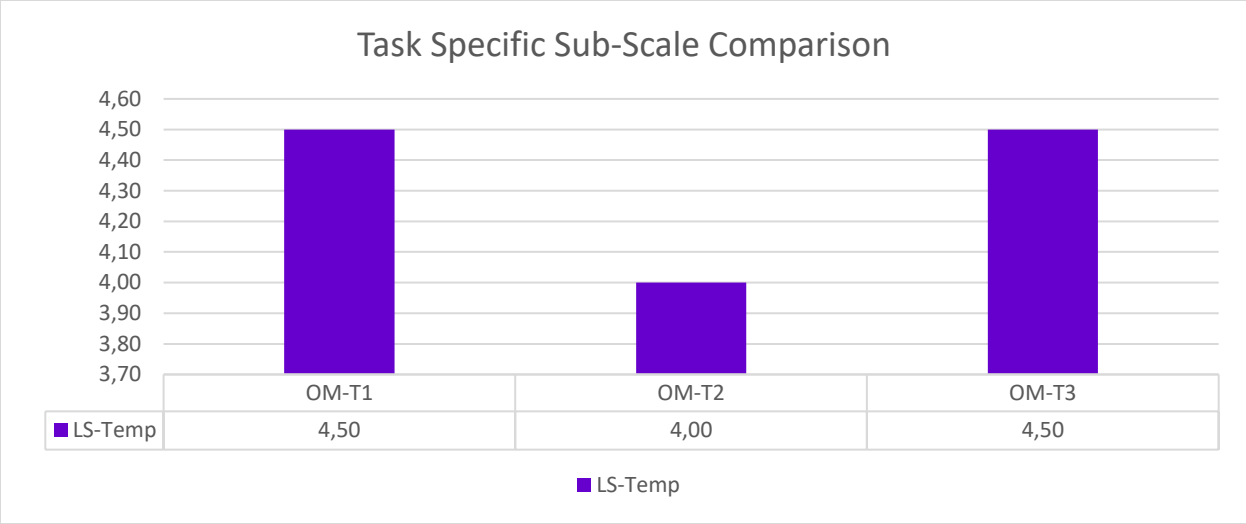
T1	T2	T3	T4	T5
Perhaps a more detailed introduction of the process and exercise at hand, could help with a faster thought process?!	NA	For practical reasons not always possible, but a mapping like this might be more fruitful and interactive in person	Same answer as step 3	Not discussed
Time boxing is not ideal for the best result	The topic felt too large to reliably prioritize and structure. An inventory alone may have been better	The network mapping would take more time and input to get to a reliable result	skipped in this meeting	skipped in this meeting
N/A	N/A	N/A	N/A	N/A
n/a	n/a	n/a	n/a	n/a

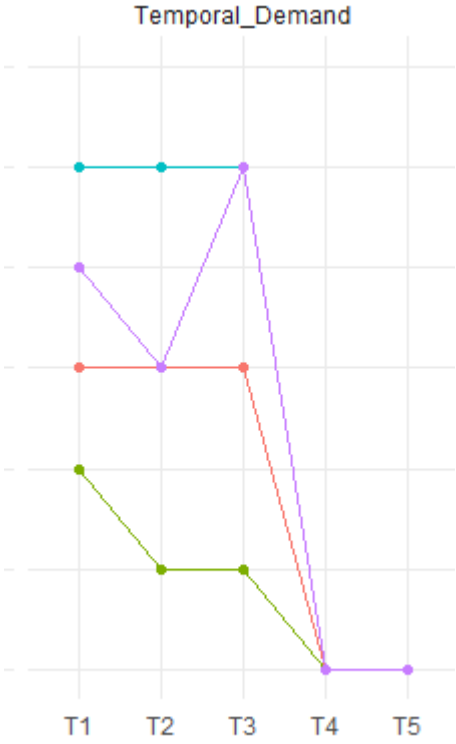
5.4.3.2 Internal Task Comparison



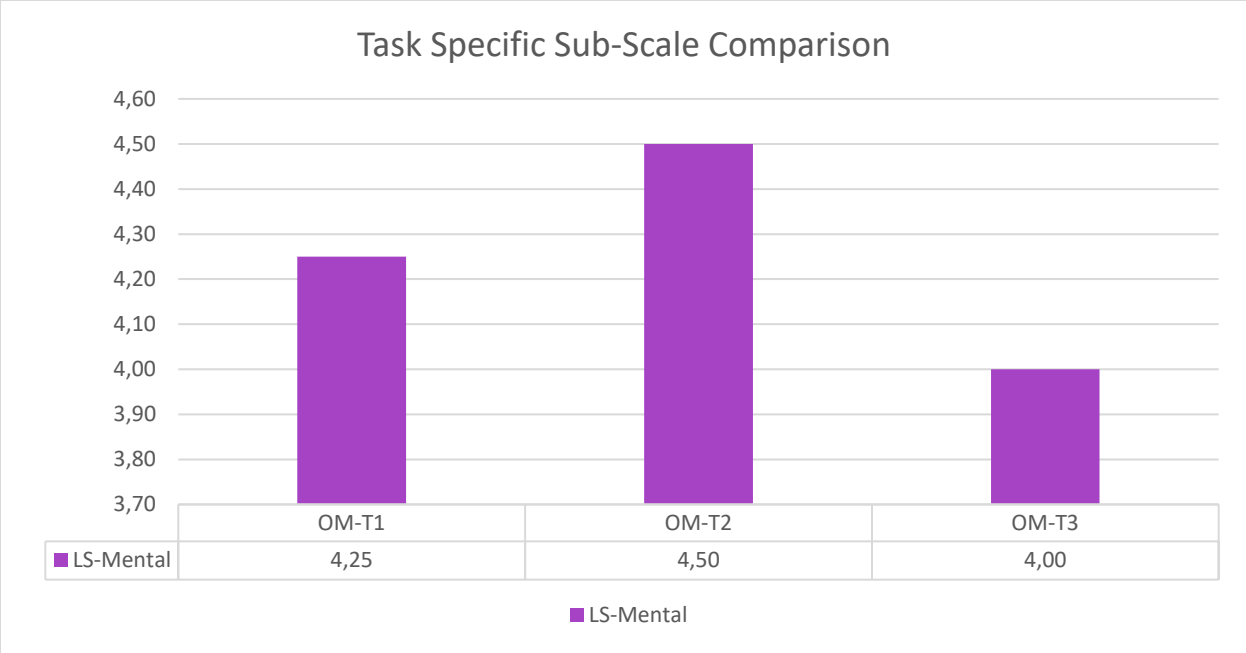


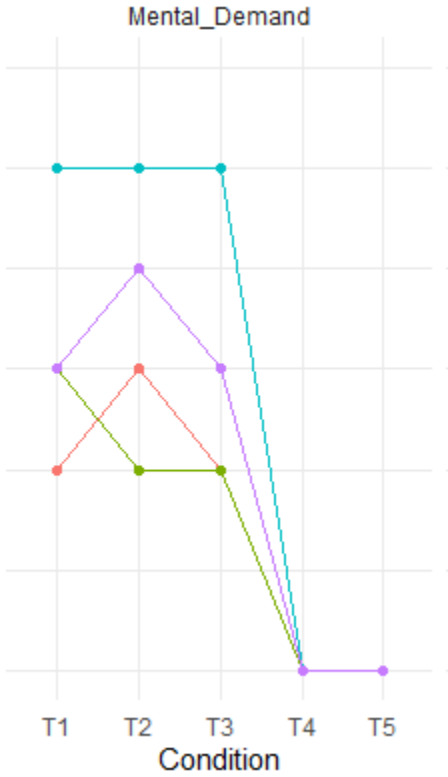
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect_ Size
Difficulty	T1	T2	4,75	4,75	0,00	1,00	0,000
Difficulty	T1	T3	4,75	4,25	0,50	0,77	-0,188
Difficulty	T2	T3	4,75	4,25	0,50	0,77	-0,188





Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect_ Size
Temporal Demand	T1	T2	4,5	4,00	0,50	0,77	-0,188
Temporal Demand	T1	T3	4,5	4,50	0,00	1,00	0,063
Temporal Demand	T2	T3	4	4,50	-0,50	0,76	0,188

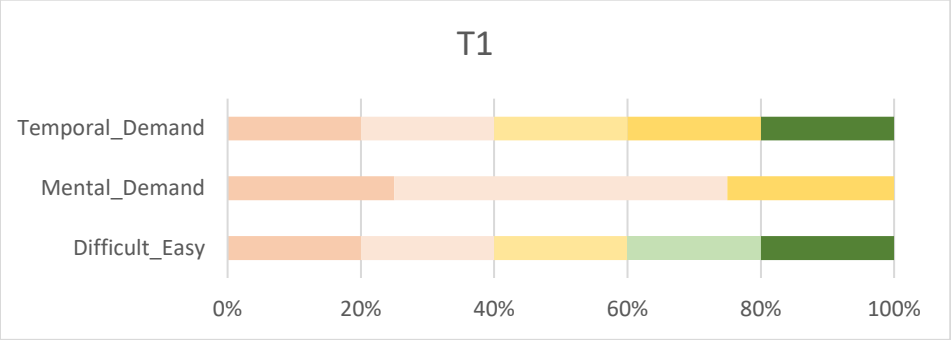




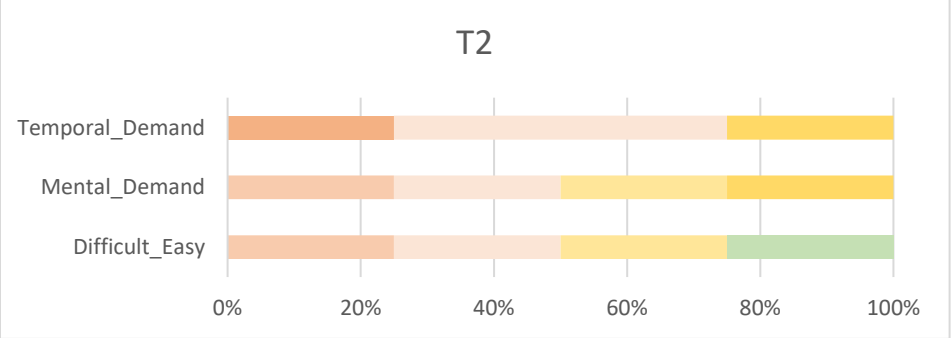
Metric	Condition 1	Condition 2	Mean Condition 1	Mean Condition 2	Mean Difference	P Value	Effect_ Size
Mental Demand	T1	T2	4,25	4,50	-0,25	0,88	0,125
Mental Demand	T1	T3	4,25	4,00	0,25	0,76	-0,188
Mental Demand	T2	T3	4,5	4,00	0,50	0,65	-0,250

5.4.3.3 Task Specific Data Distribution

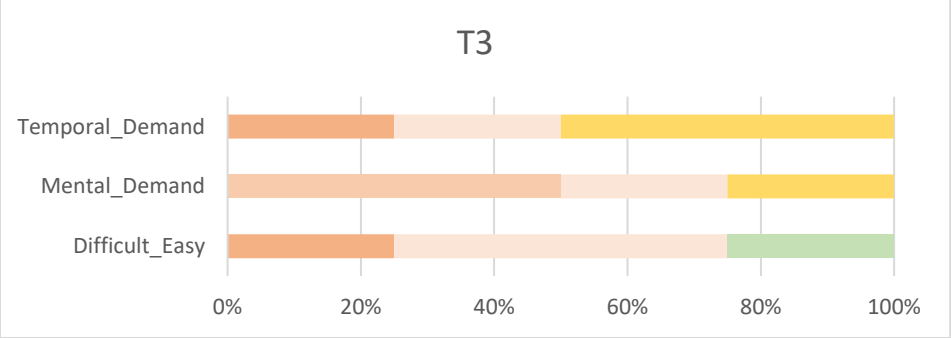
5.4.3.3.1 LS-TS – Opportunity Modelling – Task 1



5.4.3.3.2 LS-TS – Opportunity Modelling – Task 2



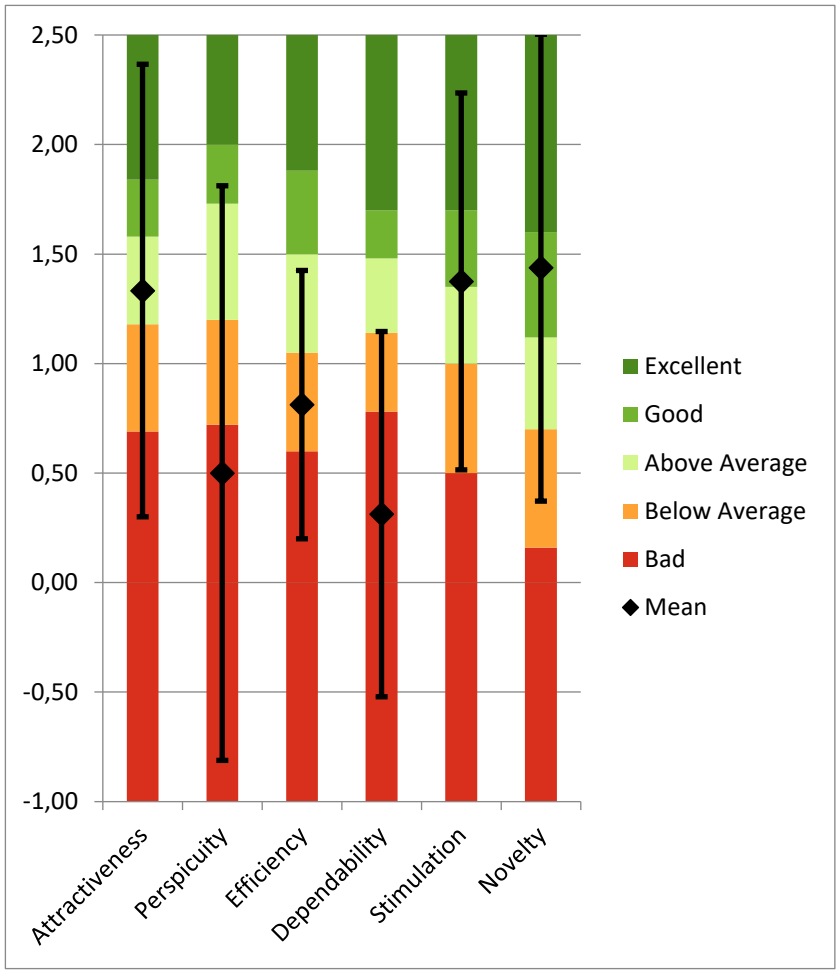
5.4.3.3.3 LS-TS – Opportunity Modelling – Task 3



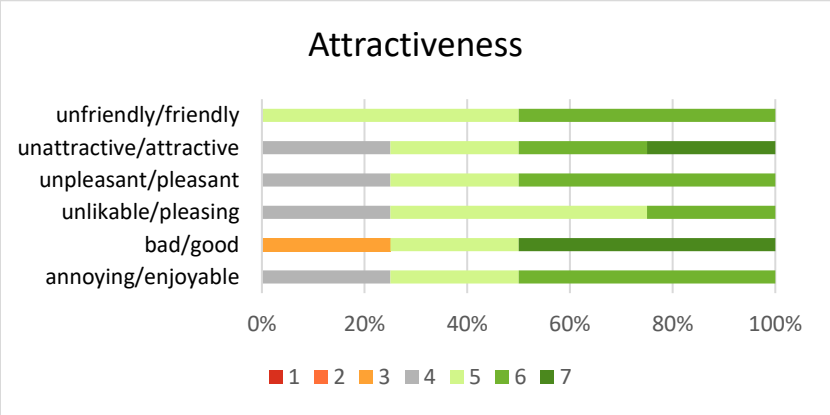
5.4.4 User Experience Questionnaire

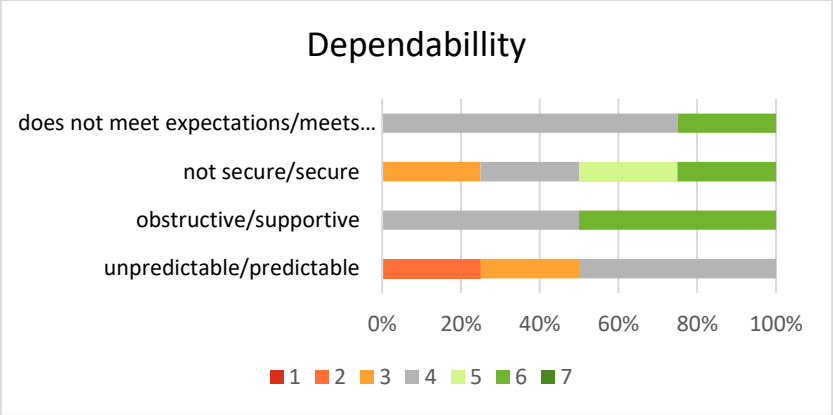
Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,33	**Above average**	25% of results better, 50% of results worse
Perspicuity	0,50	**Bad**	In the range of the 25% worst results
Efficiency	0,81	**Below Average**	50% of results better, 25% of results worse
Dependability	0,31	**Bad**	In the range of the 25% worst results
Stimulation	1,38	**Good**	10% of results better, 75% of results worse
Novelty	1,44	**Good**	10% of results better, 75% of results worse

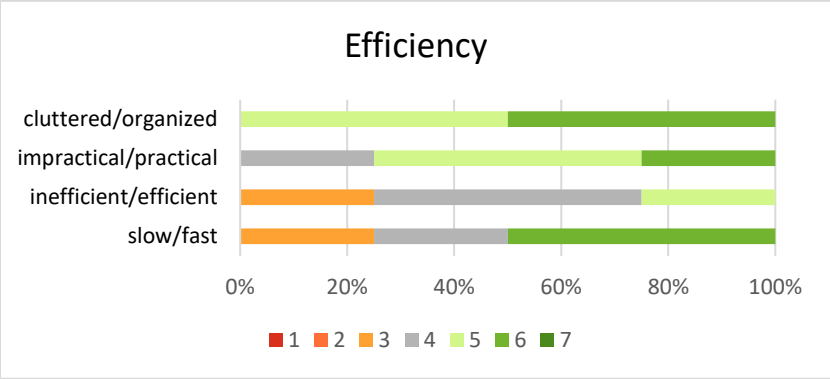
UEQ Scales (Mean and Variance)			Confidence	Cronbach Alpha
Attractiveness	1,333	1,11	1,033	0,98
Perspicuity	0,500	1,79	1,312	0,89
Efficiency	0,813	0,39	0,612	0,70
Dependability	0,313	0,72	0,834	0,74
Stimulation	1,375	0,77	0,860	0,97
Novelty	1,438	1,18	1,066	0,98

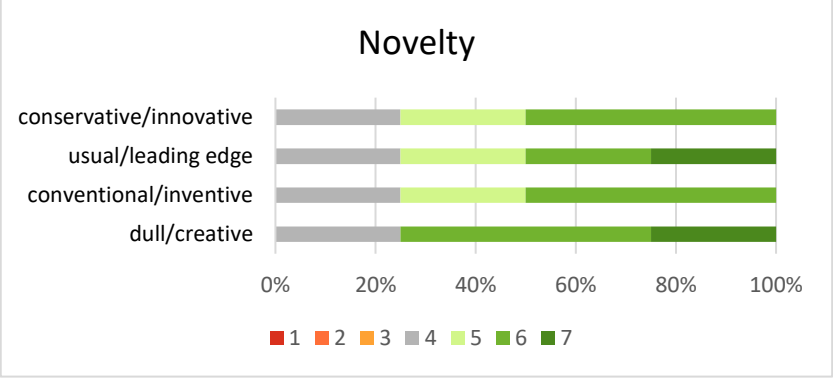


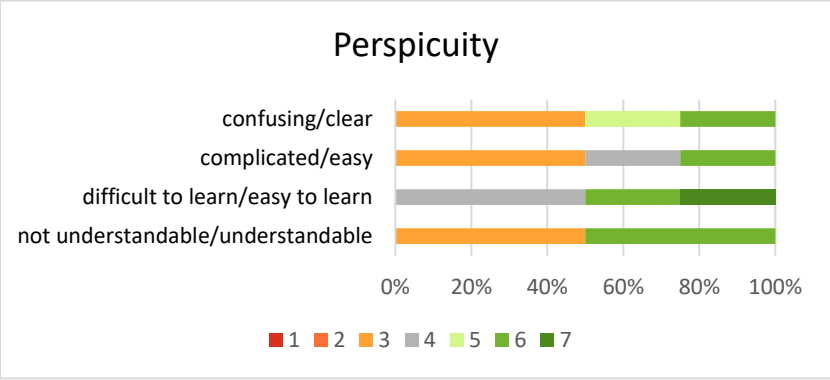
Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale
1	1,3	0,9	1,0	4	annoying	enjoyable	Attractiveness
2	0,5	3,0	1,7	4	not understandable	understandable	Perspicuity
3	1,8	1,6	1,3	4	creative	dull	Novelty
4	1,3	2,3	1,5	4	easy to learn	difficult to learn	Perspicuity
5	1,8	0,9	1,0	4	valuable	inferior	Stimulation
6	1,0	1,3	1,2	4	boring	exciting	Stimulation
7	1,5	0,3	0,6	4	not interesting	interesting	Stimulation
8	-0,8	0,9	1,0	4	unpredictable	predictable	Dependability
9	0,8	2,3	1,5	4	fast	slow	Efficiency
10	1,3	0,9	1,0	4	inventive	conventional	Novelty
11	1,0	1,3	1,2	4	obstructive	supportive	Dependability
12	1,5	3,7	1,9	4	good	bad	Attractiveness
13	0,0	2,0	1,4	4	complicated	easy	Perspicuity
14	1,0	0,7	0,8	4	unlikable	pleasing	Attractiveness
15	1,5	1,7	1,3	4	usual	leading edge	Novelty
16	1,3	0,9	1,0	4	unpleasant	pleasant	Attractiveness
17	0,5	1,7	1,3	4	secure	not secure	Dependability
18	1,3	0,9	1,0	4	motivating	demotivating	Stimulation
19	0,5	1,0	1,0	4	meets expectations	does not meet expectations	Dependability
20	0,0	0,7	0,8	4	inefficient	efficient	Efficiency
21	0,3	2,3	1,5	4	clear	confusing	Perspicuity
22	1,0	0,7	0,8	4	impractical	practical	Efficiency
23	1,5	0,3	0,6	4	organized	cluttered	Efficiency
24	1,5	1,7	1,3	4	attractive	unattractive	Attractiveness
25	1,5	0,3	0,6	4	friendly	unfriendly	Attractiveness
26	1,3	0,9	1,0	4	conservative	innovative	Novelty

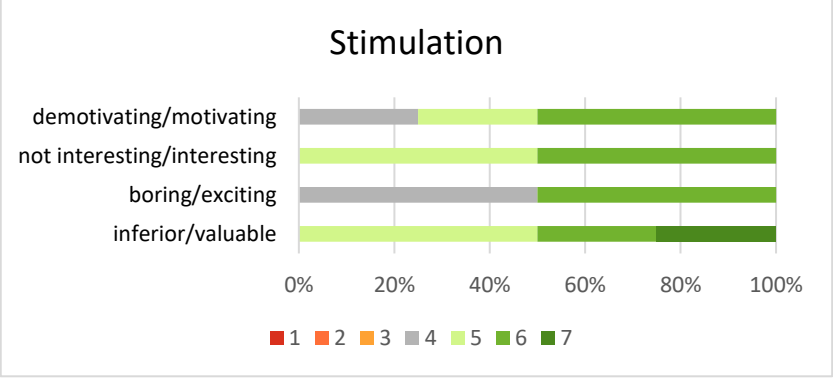

Attractiveness
unfriendly/friendly
unattractive/attractive
unpleasant/pleasant
unlikable/pleasing
bad/good
annoying/enjoyable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Dependabillity
does not meet expectations/meets...
not secure/secure
obstructive/supportive
unpredictable/predictable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Efficiency
cluttered/organized
impractical/practical
inefficient/efficient
slow/fast
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Novelty
conservative/innovative
usual/leading edge
conventional/inventive
dull/creative
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Perspicuity
confusing/clear
complicated/easy
difficult to learn/easy to learn
not understandable/understandable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7


Stimulation
demotivating/motivating
not interesting/interesting
boring/exciting
inferior/valuable
0%
20%
40%
60%
80%
100%
1 2 3 4 5 6 7

5.5 SCP – Intrigue and Perceived Value