Pontifícia Universidade Católica
do Rio de Janeiro

Bianca Rodrigues Teixeira

Investigating the integration of user values with design rationale and its effects on HCI design artifacts

Dissertation Proposal

Dissertation presented to the Programa de Pós–graduação em Informática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática.

Advisor: Profª Simone Diniz Junqueira Barbosa

Rio de Janeiro
August 2020

Pontifícia Universidade Católica
do Rio de Janeiro

Bianca Rodrigues Teixeira

Investigating the integration of user values with design rationale and its effects on HCI design artifacts

Dissertation presented to the Programa de Pós–graduação em Informática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática. Approved by the Examination Committee.

Profª Simone Diniz Junqueira Barbosa
Advisor
Departamento de Informática – PUC-Rio

Profª Carla Faria Leitão
Departamento de Psicologia – PUC-Rio

Drª Juliana Soares Jansen Ferreira
IBM – Brasil

Rio de Janeiro, August 18th, 2020

Bianca Rodrigues Teixeira

Bachelor in Information Systems (2017) at the Federal University of the State of Rio de Janeiro (UNIRIO). Since 2019 she works at ExACTa at PUC-Rio.

Bibliographic data

Teixeira, Bianca Rodrigues

Investigating the integration of user values with design rationale and its effects on HCI design artifacts / Bianca Rodrigues Teixeira; advisor: Simone Diniz Junqueira Barbosa. – Rio de Janeiro: PUC-Rio, Departamento de Informática, 2020.

v., 83 f: il. color. ; 30 cm

Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui bibliografia

1. Valores de usuário. 2. Design rationale. 3. Design Sensível a Valores. 4. Questões Opções e Critérios - QOC. 5. Linguagem de Modelagem para Interação como Conversa - MoLIC. I. Barbosa, Simone Diniz Junqueira. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Acknowledgments

I would like to thank my advisor, Simone, for always being extremely patient and thorough. Despite being busy, she still managed to find time to help me and I truly appreciate her assistance and advice.

Thank you to the jury members, Carla and Juliana, for the availability and for being so helpful and for providing great insights. It feels great to have these three female role models in the Computer Science field to look up to, along with Clarisse, who recently retired but remains an inspiration.

I would also like to thank everyone who participated in the studies we conducted. I am grateful to have such proactive colleagues who are always willing to help out their peers. It feels great to have been a part of this group.

Lastly, I would like to thank my family and friends for always supporting me. I love you all.

This study was financed in part by the National Council for Scientific and Technological Development - Brasil (CNPq) - Finance Code 001

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Teixeira, Bianca Rodrigues; Barbosa, Simone Diniz Junqueira (Advisor). **Investigating the integration of user values with design rationale and its effects on HCI design artifacts**. Rio de Janeiro, 2020. 83p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

User values are a significant part of designing software, but are not always considered explicitly in the design process. When making design decisions, user values can get lost among the designers' own biases regarding their values. To avoid this pitfall, this work studies the integration of user values with design rationale techniques, namely Questions, Options, and Criteria (QOC), and how this integration reflects on a resulting design artifact (in particular, an interaction diagram using MoLIC). We performed two separate studies with Computer Science students in Rio de Janeiro. In the first study, we found that user values can be integrated into design rationale using informal notations, such as circling or underlining the options or criteria related to user values. The decisions made with the QOC method using user values did result in relevant impacts on MoLIC diagrams. Participants who performed activities for users with strong values had richer results than those for more "generic" users. In a second study, we found that designers can recognize when user values are embedded into design artifacts. These results are encouraging to continue research regarding user values, with possibilities of developing new methods or updating existing techniques and notations such as QOC or MoLIC to explicitly support user values.

Keywords

User values; Design rationale; Value Sensitive Design; Questions Options and Criteria - QOC; Modeling Language for Interaction as Conversation - MoLIC.

Resumo

Teixeira, Bianca Rodrigues; Barbosa, Simone Diniz Junqueira. **Investigando a integração de valores de usuário com design rationale e seus efeitos em artefatos de design de IHC**. Rio de Janeiro, 2020. 83p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Valores de usuário são um fator importante no design de software, mas nem sempre são considerados de forma explícita no processo de design. Ao tomar decisões de design, valores de usuários podem se perder junto aos viéses dos designers relacionados a seus próprios valores. Para evitar esse problema, este trabalho estuda a integração de valores de usuários com técnicas de *design rationale*, especificamente Questões, Opções e Critérios (QOC), e como essa integração é refletida em um artefato de design resultante (um diagrama de interação usando MoLIC). Conduzimos dois estudos separados com estudantes de Ciência da Computação no Rio de Janeiro. No primeiro estudo, vimos que valores de usuários podem ser integrados com *design rationale* usando notações informais, como circulando ou sublinhando as opções ou critérios associados a valores de usuários. As decisões tomadas com o método QOC usando valores de usuário resultaram em impactos relevantes nos diagramas MoLIC. Participantes que realizaram as atividades para usuários com valores bem definidos tiveram resultados mais ricos do que aqueles com usuários mais "genéricos". No segundo estudo, vimos que designers conseguem reconhecer quando valores de usuários são incorporados em artefatos de design. Esses resultados são estimulantes para continuar a pesquisar valores de usuários, com possibilidades de desenvolver novos métodos ou de atualizar técnicas e notações existentes, como QOC ou MoLIC, para dar suporte a valores de usuários.

Palavras-chave

Valores de usuário; Design rationale; Design Sensível a Valores; Questões Opções e Critérios - QOC; Linguagem de Modelagem para Interação como Conversa - MoLIC.

Table of contents

List of figures

List of tables

List of abbreviations

DR – Design Rationale
HCI – Human-Computer Interaction
IBIS – Issue-Based Information System
MoLIC – Modeling Language for Interaction as Conversation
RQ – Research Question
SE – Software Engineering
SemEng – Semiotic Engineering
QOC – Questions, Options, and Criteria
UCD – User-Centered Design
UI – User Interface
VSD – Value Sensitive Design

1
Introduction

In the field of Human-Computer Interaction (HCI), in most approaches the user is considered the focus of interactive systems design. Therefore, the users' goals, preferences, needs, and, most importantly, values should be investigated in order for designers and developers to deliver appropriate solutions. As most research has investigated goals, needs, and preferences (Carroll, 1997; Fischer, 2001; de Souza, 2005b), in this work we focus mainly on human values.

After performing a literature review, Schwartz and Bilsky (1987) state that values are generally defined as "*(a) concepts or beliefs, (b) about desirable end states or behaviors, (c) that transcend specific situations, (d) guide selection or evaluation of behavior and events, and (e) are ordered by relative importance.*" Based on this definition, they provide the following interpretation for human values: "*Values are cognitive representations of three types of universal human requirements: biologically based needs of the organism, social interactional requirements for interpersonal coordination, and social institutional demands for group welfare and survival*" (Schwartz and Bilsky, 1987).

As values guide selection or evaluation of behavior and events, they can also be defined, as Tisdale (1961) proposed, as motivational constructs. In his Ph.D. thesis regarding psychological value theory, Tisdale (1961) provides the following definition: "*Values are inferred motivational constructs associated with perceived differences in goal-directed behavior and indicated by the selection of action alternatives within social situations*". Among the definitions found in the literature, we consider this one as the most appropriate for our work. The idea of having a goal-directed behavior in which a person must pick between alternatives will be used often in this research, as we will describe further in this section, and we believe that values play a big part in these decisions, especially in a software design context.

Kujala and Väänänen-Vainio-Mattila (2009) also conducted a literature review and defined user values as being "users' internal conceptions of what is important in a certain usage context and they are not perceptions of products". In addition, they relate user values to software development, in the sense of the impact user values can have on software development.

The concept of user values makes the motivational aspect of

> system/product usage visible to developers. The values represent both users' preferences as to what is important to them and aversions to what they want to avoid. (Kujala and Väänänen-Vainio-Mattila, 2009)

In order to incorporate human values into the design of information and computational systems, Friedman et al. (2008) developed *Value Sensitive Design* (VSD), defining it as a framework for technology design that accounts for human values throughout the design process in a "principled and comprehensive manner". Value Sensitive Design features a *tripartite methodology*, which includes conceptual, empirical, and technical investigations (Friedman et al., 2008). Conceptual investigations aim at answering questions such as, "*What stakeholders are directly and indirectly affected by the design?*". It is about gathering conceptual aspects of the context in which the design is situated. Empirical investigations, on a different note, answer questions based on observation, measurement or documentation of any human activity. They can include surveys, interviews, and other methods. Finally, technical investigations focus on the impact of existing technologies on human values. They also involve designing systems that take into account values identified in conceptual investigations. We discuss VSD further in Section 2.1.

When designing a solution, many decisions need to be made in a wide spectrum of issues, including the goals that will be supported, the task structure and user-system interaction to support them, the concrete user interface layout and elements, to name a few. For instance, a software company may have to choose one feature over another, or decide between two colors for the application logo. As many alternative solutions can be considered, reflecting upon all the alternatives and keeping track of the underlying reasons for selecting one solution over another can be a daunting task (Carroll and Rosson, 2003). As a way to support and document this kind of decision making, design rationale methods can be adopted. Design rationale can be defined as a record of the decisions that led to a specific design choice, which can then be reflected on an artifact or on a feature (Lee and Lai, 1991). Several design rationale techniques and notations have been proposed. This will be further discussed in Section 2.2.

We argue that, when dealing with values and value tensions (*i.e.*, conflicts between values), recording the design rationale becomes essential. However, to the best of our knowledge, the integration of Value Sensitive Design, or user values, and design rationale has not been explored. The studies found in the VSD literature usually relate the final product to the values used to help define features or software requirements, but how this relation is

established is unclear. The decision-making process can easily get lost and be forgotten. Friedman et al. (2008) state that "*Value Sensitive Design seeks to be proactive to influence the design of technology early in and throughout the design process.*" This work aims to study this influence with the following research questions:

RQ1: How can user values be integrated into design rationale?
RQ2: How can design artifacts reflect the design rationale taking into account user values? – RQ2 can be unfolded in two subquestions:

RQ2a: Given a design rationale which explicitly references values, do designers embed those values in the design artifacts, explicitly or implicitly?
RQ2b: If a design artifact embeds values, will designers recognize those values?

Since we are interested in studying design artifacts and how they can communicate user values, this work will be grounded in Semiotic Engineering (SemEng). Proposed by de Souza (2005b), this theory brings the designers to the front stage, prompting them to reflect upon and make explicit not only their understanding of the users and their tasks, but all their (the designers') assumptions and decisions that inform the design. In RQ2a, we study the emission of user values, whereas in RQ2b, we focus on the communicability and how values are received and interpreted by designers. Therefore, as a way to answer RQ2, the design artifact to be used will be diagrams written in MoLIC, an interaction modeling language based on semiotic engineering.

This document is organized as follows. Chapter 2 describes theoretical foundations and related work regarding design rationale, user values, Semiotic Engineering, and MoLIC. In order to evaluate the effects of integrating user values into Design Rationale on design artifacts, we conducted two studies, as described in Chapter 3. In Chapter 4, we go through the preparation and the results obtained from Study 1 and in Chapter 5 we discuss the preparation and results from Study 2. Finally, in Chapter 6, we conclude this work.

2
Theoretical foundations and related work

In this chapter, we describe the background research and related work. In Section 2.1, we describe user values. In Section 2.2, we describe the principles of design rationale and some techniques used to document it. In Section 2.3, we describe the main concepts of Semiotic Engineering related to this work. In Section 2.4, we describe Interaction Modeling with the MoLIC language.

2.1
User values

Davis and Nathan (2015) discuss applications, adaptations, and critiques of Value Sensitive Design (VSD) in the literature. Among their findings, they state that the field of VSD is far from ideal, since the design for human values is not yet considered routine. They compare VSD with User-Centered Design (UCD), a set of principles and activities widely adopted today in the field of HCI (Davis and Nathan, 2015). The authors claim that VSD can be the next UCD, in that sense.

> "Work in the area will continue to grow more nuanced and more reflective. Focusing design on human values will become an accepted rather than novel perspective. Attention to the user is infused throughout HCI work and gaining ground in software development practice, even when there is no explicit reference to UCD; we hope that someday attention to values will be just as pervasive, even if VSD (or another branded values-oriented methodology) is rarely referred to. What we learn from engaging with VSD today will influence how technology designers appreciate and address values in the future." (Davis and Nathan, 2015)

Friedman and Kahn (2002) state that it is imperative to take a proactive stance in human values and ethics along with other multidisciplinary collaborations. According to the authors, it is much easier to design systems with values in mind from the beginning than to revisit poor systems that have already been firmly established and rooted within organizations. Human values should be considered as a design criterion with an ethical import – similarly to

how reliability, correctness, and efficiency are traditional software criteria, "compliance" with user values should also be crucial (Friedman and Kahn, 2002).

Friedman et al. (2017) conducted a survey on VSD methods. It includes *direct and indirect stakeholder analysis*, which identifies those affected by the technology under investigation, either directly or indirectly, while also listing how they may be affected by benefits, harms, and tensions. It is important to find what specific roles need to be represented in the analysis, which is performed after interviews with participants in the context of the study (Denning et al., 2014).

Value-oriented semi-structured interview aims at understanding the stakeholders' views and values about a technology, while also exploring value tensions. The semi-structured aspect of the interview allows for some topics to be examined more deeply but still leaves room for the interviewer to engage with new considerations brought by the stakeholder (Friedman et al., 2017). The interviews can feature audio recording and transcriptions for analysis (Friedman, 1997; Friedman et al., 2009). Different tactics can be employed depending on the profile of the stakeholders involved. For instance, when interviewing children about robotic pets, Kahn Jr. et al. (2006) let them hold the artifacts in question, which helped the children remain focused on the topic.

Value source analysis involves identifying and distinguishing the official project values, values held by designers and values held by stakeholders. This distinction is important to make sure that the final project does not feature too many strong values held by only designers and/or stakeholders, and that it maintains its core values. These core values are described as "explicitly supported project values" and are agreed upon to guide the design and development process (Friedman et al., 2017).

Value scenarios and *Value sketches* can act as a values representation and as a values elicitation method (Friedman et al., 2017). The former comprise textual narratives that emphasize direct and indirect implications for the stakeholders, key values, and effects of widespread and long term use of the technology. They can be written by the researchers or the participants of the study. The latter aim at providing understandings, views and values about a technology, through drawings. Participants can show the researchers, in a visual way, what is important to them in regard to the technology in question.

The *Value Dams and Flows* method works by defining value dams, *i.e.*, undesirable features, and value flows, *i.e.*, desirable features (Friedman et al., 2017; Miller et al., 2007). It works by "*(a) avoiding features that even a small number of stakeholders view as particularly problematic, (b) identifying*

and designing for values stakeholders wish to see the system embody, and (c) systematically addressing values-oriented design tradeoffs." (Miller et al., 2007).

When using this method, the designer must balance the value dams and flows, which can generate conflict and could go against some users' preferences. As the authors state, "*When value tensions go unaddressed, consequences can range from lack of appropriation by disadvantaged groups to more severe consequences such as system sabotage.*" (Miller et al., 2007). Before applying the method, the authors conducted a survey to define harms and benefits as dams and flows, as perceived by stakeholders. Each harm and benefit statement was related to a specific value, such as "privacy" and "reputation". The authors used a threshold of 50% rate of agreement with "benefit" statements to establish value flows and a 10% rate of agreement with "harms" statements to establish the dams. An example of a privacy harm statement considered a value dam is: "I would feel like my privacy is being compromised if the system logged what and how I searched".

The authors conclude, in their case study, that the Value Dams and Flows method helped mitigate some value tensions, but not all. The users of the system, which was then implemented with the help of the method, in a survey, later claimed to be concerned with their reputation, whereas only 1 out of 6 reputation-related statements was identified as a value dam, thus not given much attention in the design and development process. The authors conclude by suggesting the implementation of a feature to remedy this problem in the future.

An issue we find with the Value Dams and Flow Method is that the statements provided by the survey are swayed to one side. They are either classified as a harm or as a benefit. Yet, many statements can have more than one side. For instance, the privacy statement disclosed above could be rephrased as: "I would like to see my history of searches". This becomes a benefit related to the *transparency* value, and the stakeholders could deem it as a positive thing, thus, as a value flow. The feature *per se* is the same – system logs. But the phrasing can change the users' perception and consequently change whether it is a desirable or undesirable feature, *i.e.* a value flow or a value dam.

Also, we believe Friedman et al.'s definition of human values is simplistic and use overly broad terms. They define *value* as what a person or group of people considers important in life (Friedman et al., 2008). In the context of technology design, we believe this interpretation can wrongfully consider software features as being human values. As seen in the definition of the Values

Dams and Flows method, the concepts of *value* and *feature* seem to be used interchangeably. For example, the feature that logs user searches was deemed as a value dam in Miller et al.'s study. The value in question, however, is privacy, not system logs. The authors use system features, which are related to specific values, as the input to the Value Dams and Flows method, which seems counter intuitive. To avoid confusion with software features, this has motivated us to adopt the definition of human values proposed by Tisdale (1961) and presented in Chapter 1, which is: "*Values are inferred motivational constructs associated with perceived differences in goal-directed behavior and indicated by the selection of action alternatives within social situations*".

Borning et al. (2005), using Value Sensitive Design methods, developed an application (UrbanSim) that simulates urban development to inform the public when deliberating land use and transportation decisions. Their work centered on public deliberation and decision making involving multiple stakeholders, and they used simulation to convey long-term consequences of alternative choices to *inform* the decision-making process. In doing that, the authors categorized the indicators of urban planning into three broad value categories (economic, environmental, and social). The stakeholders, then, were able to navigate the application and make better informed decisions according to their personal values. This work, however, did not structure or document the underlying design rationale.

Ferrario et al. (2016) introduced the concept of Values-First Software Engineering. It explicitly uses human values in decision-making processes during key stages of software development, and systematically maps all values in a project independently of moral judgement. According to the authors, Values-First SE is different from VSD in the sense that it is more practical and less "grounded in theory". Values-First SE creates a "values taxonomy" (*i.e.*, universal values), which is reused throughout the process of decision making, and VSD allows for a more open interpretation to the dimension of values. Also, VSD favors values with ethical components, whereas Values-First SE gives equal representation to all human values (Ferrario et al., 2016).

They also performed a study to investigate and measure how human values influence software production decision-making processes, using Schwartz's universal values model (Winter et al., 2018; Schwartz and Bilsky, 1987). They used Values Q-Sort, a technique in which participants of the study were asked to sort statements by level of agreement while also being interviewed. The authors asked the participants to focus the sorting process on a specific software project, to help anchor the reasoning (Winter et al., 2018). They found three profiles of software engineers: the "Intrinsically-driven, Socially-concerned Soft-

ware Engineer", the "Autonomous, Nonconforming Risk-taker", and the "Fun-loving, Extrinsically-driven Software Engineer". Their work of understanding software engineers and their general behavior contributes greatly to the SE community, but, with an HCI perspective, the software users' values should also be studied. The possible bias of a software engineer's values in the software he creates can be an issue for the final user and her values. This work aims at studying how designers take into account users' values, while also dealing with possible conflicts regarding the designers' own values.

Winter et al. (2019b) have been advancing their work regarding values in software engineering by developing different methods. The Values Survey method collects qualitative data on the relationships between values. The Values Probes method was designed to provoke more disruptive and creative thinking among software engineers. Along with the Values Q-Sort, which combines both a qualitative and a quantitative approach, the Values Probes and Values Survey methods help support the study of values within the SE community (Winter et al., 2019a,b).

Harbers et al. (2015) conducted a study using their "Value Story workshop", which joins the elicitation of user values with user stories, a common format for eliciting requirements in the Requirements Engineering field. The technique comprises five steps: (1) analysis of system stakeholders; (2) analysis of system stakeholders' values; (3) list of concrete situations with the values; (4) identification of stakeholders ' needs for each concrete situation; and (5) creation of user stories (Detweiler and Harbers, 2014). In their study, the authors compared user stories created through Value Story (value-based user stories) and through regular requirement elicitation practices (regular user stories).

A regular user story has the following template: "As a <role>, I want <something> so that <benefit> (where the last part of the user story (so that <benefit>) is optional)". A value-based user story has a slightly different template: "As a <stakeholder>, I want <stakeholder need> in order to support <value>" (Harbers et al., 2015).

In the Value Story workshop conducted, the first step – stakeholder analysis – was not completed due to time constraints. An identification of stakeholders can be an arduous task, and including it as step one of a user story elicitation workshop may not result in a thorough enough analysis. If a persona had been previously created with a specific care for user values, perhaps it would have been possible to carry on the workshop with a complete analysis of all stakeholders' needs, replacing steps (1) and (2) with an analysis of the personas.

The resulting value-based user stories and regular user stories were ana-

lyzed by both software developers and VSD experts. Because of their different perspectives on user stories, each group analyzed specific factors. Developers analyzed each user story according to the INVEST criteria (Independent, Negotiable, Valuable, Estimable, Small, Testable), whereas VSD experts were asked to identify up to three values for each user story – which can be either hindered, supported or unaffected by the story –, as well as to provide scores for: (1) how much the developer will understand how the desired feature will affect the value(s) at stake; and (2) how much the value perspective is explicitly addressed in this user story (Harbers et al., 2015).

Overall, regular user stories had higher scores when analyzed by software developers, and value-based user stories had higher scores when analyzed by VSD experts. According to the developers, value-based user stories fell short in being "estimable", "small" and "testable", and according to the VSD experts, regular user stories were not understandable regarding how they will affect user values and also not explicit regarding the value perspective. These results leave room for improvement on methods for designing systems with a focus on user values.

When it comes to software code, Mougouei (2020) proposed a notion of "Value Programming", which is composed of three principles: (P1) Value Annotation: specifying the relevance of code elements (e.g., classes and methods) to human values; (P2) Value Inspection: inspecting source code to detect conditions that lead to potential value breaches in software (value smells); and (P3) Value Recommendation: making recommendations to address values and mitigate value breaches and biases in software. Value annotation can be done on APIs with machine learning techniques, and on software code, either manually or automatically – parts of the code that interact with the annotated APIs consequently will feature the same values (Mougouei, 2020). Value inspections can identify conditions/faults that breach annotated values, and the value recommendation component aims at mitigating these conditions, which can happen in a high or low level. Although still in early stages, this framework is interesting as it relies on software developers to consider and remember user values when writing code. Value annotations have the potential of being a constant reminder of user values, and the inspections and recommendations can ultimately result in software products that respects the users' values more.

Thew and Sutcliffe (2008) proposed a taxonomy of user values for the Requirement Engineering process, which includes nine values, their potential sources, and their process implications – for instance, "morals/ethics" can be assessed by analyzing "behaviour towards others and opinions of others' behaviours", and can result in "open process, use of workshops to promote

inclusivity". The "process implication" aspect can impact how the development team is organized, as the taxonomy focuses on the RE process, not necessarily on the resulting elicited requirements.

As Sutcliffe (2013) claims, user values may have implications for non-functional requirements and impact directly design decisions concerning functional requirements. The author also provided another, similar taxonomy of seven user values, each with its own generic functional requirements, design implications/rationale and arguments. For example, "morals/ethics" can result in "visibility and transparency and data access, shared controls", which can be achieved by designing "behavior monitors and configurable controls" as it is "closely related to trust". This taxonomy is supposed to be "generic", but we believe that user values are directly related with a specific context. Since, according to Tisdale (1961), values are indicated by the selection of alternatives within social situations, different systems or software will pose different situations.

Other values can be elicited using the Value-Based Requirements Engineering method, which uses scenarios and the taxonomy proposed in (Thew and Sutcliffe, 2008) to guide interviews focused on values (Sutcliffe, 2013).

Mougouei et al. (2018) provided a research roadmap for operationalizing human values into software, with a Software Engineering perspective, which focuses on (1) establishing practical definitions for human values, (2) integrating values into software design, and (3) measuring values in the software development life cycle. They enumerated 21 research questions to be investigated by researchers, including: "(RQ3): How do software artifacts embed human values?", "(RQ7): How do developer values impact software?", "(RQ10): How do design practices embed or breach values?", "(RQ13): How can existing design practices be extended to account for human values?". These research questions are closely related to this work, which is encouraging and helps validate our motivation.

2.2 Design rationale

Several approaches on design rationale have been proposed, and a few surveys have been conducted. We summarize here relevant findings from the surveys by Yue et al. (2018), Jarczyk et al. (1992), and Regli et al. (2000).

Kunz and Rittel (1970) proposed IBIS (Issue Based Information Systems) as a means to support coordination and planning of political decision process. This work is considered to be the pioneer in the field of design rationale, even though the authors did not use the term *design rationale* to describe it. It

focuses on solving "issues", *i.e.*, controversial questions that are first raised, then argued, settled, dodged, or substituted. IBIS was also applied in design processes (Noble and Rittel, 1988).

gIBIS (or graphical IBIS) is a hypertext-based, graphical implementation of the IBIS method, created to facilitate the capture of early design deliberations (Conklin and Begeman, 1987). It captures the design history – decisions, rejected options and trade-off analysis – and supports multiuser usage. Many users found the IBIS method powerful for research thinking and design deliberation (Conklin and Begeman, 1988).

Another relevant work is the Questions, Options, and Criteria (QOC) approach, by MacLean et al. (1991). It is simple, and comprises *questions*, which identify key design issues, *options*, *i.e.*, possible answers to the questions, and *criteria*, used to assess and compare the options. This assessment can be either positive or negative, and contribute to an overall scenario that will ultimately lead to an informed decision to select an option for the question. The QOC notation simplifies the possibilities and the rationale behind each option.

Tang et al. (2006) performed an interesting study with designers and software architects on their perception and usage of design rationale. They found that practitioners often forget the reasons that justify design decisions, and that it is difficult to understand design by other designers without the design rationale that supports it. The respondents also rated nine common reasons for making software architecture-related decisions as being extremely important, such as decisions based on the benefit of the design. According to their study, almost 80% of professionals frequently document design rationale motivated by design constraints and assumptions, which the authors found encouraging. However, when architects and designers do not document design rationale, they do so for time or budget reasons, or by not having specific standards or a suitable tool.

To the best of our knowledge, there are no design rationale methods that explicitly incorporate user values. We believe this could be done by providing a checklist of user values, collected by methods described in Section 2.1, such as Values Survey, to be considered in every decision, as a reminder for decision makers. Perhaps, for every option or alternative, the decision maker could document what user value motivated the decision to adopt or abandon each option.

2.3 Semiotic Engineering

Semiotic Engineering (SemEng) assigns to both users and designers the role of interlocutors in a communicative process (de Souza, 2005b). Designers convey their intentions about what they expect the users to understand and to achieve, and about how the users, in turn, are expected to respond. This happens via the user interface (UI), through words, images, and other signs. The UI, which is created by the designer for the user, is viewed as the *designer's deputy*. SemEng makes the designers aware of their role in the HCI process, as a communication process, which is to prepare the ground for the experience the user will have with the system. The interaction shall happen according to the way the designer previously anticipated.

The designer, then, is considered a first-class citizen. By communicating with users through a specifically designed space at interaction time, the designers' role as active interlocutors demonstrate the importance of bringing users and designers together at interaction time (de Souza, 2005a). Thus, software artifacts, created as a result of the designer's human reasoning, should be crafted with care, avoiding communicative breakdowns. It is the designer's job to consider a myriad of possibilities of system usage, so that the system interface, or the designer's deputy, can provide an adequate response. By understanding the user and his/her values, this process becomes better informed.

Semiotic Engineering characterizes interactive systems as metacommunication artifacts, which communicate a message about communication itself (de Souza, 2005b). This metacommunication message can be paraphrased as:

> Here is my understanding of who you are, what I've learned you want or need to do, in which preferred ways, and why. This is the system that I have therefore designed for you, and this is the way you can or should use it in order to fulfull a range of purposes that fall within this vision. (de Souza, 2005b)

The designer's messages trigger a semiosis process in the users, who interpret and decode them. These messages comprise signs, which are representations intended by the designer to convey certain meanings to users. A sign can be an image, a word or text, a widget, etc. Ideally, the meaning interpreted by the users would be the same as that intended by the designers, so that communication can happen smoothly. Since this is not always the case, and the designers will not be present at interaction time to negotiate those meanings

with the user, the designers must be careful when conceiving and expressing the signs with which the users will interact.

SemEng has two methods for evaluating the metacommunication: the semiotic inspection method and the communicability evaluation method (de Souza and Leitão, 2009).

The semiotic inspection method (SIM) examines the different kinds of signs available on a user interface and product documentation. Signs can be static, *i.e.*, signs that convey a message that can be understood at a glance, by "reading" the user interface; dynamic, *i.e.*, signs that require interaction to be fully understood, which may take longer; and metalinguistic, *i.e.*, signs that refer to other signs (*e.g.*, documentation and help content). SIM also includes a contrasting comparison of the metacommunication messages sent from designer to user obtained from each inspection. It concludes with an appreciation of the quality of the overall designer-to-user metacommunication (de Souza et al., 2006). It focuses on the emission of the message, similarly to RQ2a, in which we are interested in whether user values are apparent in design artifacts.

By contrast, the communicability evaluation method (CEM) helps designers to grasp how well the users are getting the intended messages during interaction and to identify communication breakdowns, *i.e.*, when the interaction does not go as expected and fails. Since users are only communicating with the designer's deputy and not the designer herself, it is important to learn whether the designer's conveyed messages are being well received and understood. During the execution of the method, the users can express their impressions on the communication messages and HCI design choices (Prates et al., 2000). In contrast to SIM, CEM focuses on the reception of the message – similarly to RQ2b, in which we study whether designers recognize user values in design artifacts. CEM has three major steps: tagging, interpretation, and semiotic profiling. The first identifies points with communication breakdowns, the second maps these breakdowns onto HCI problems – done by an HCI expert –, and the third – done by a SemEng expert – provides a characterization of the overall message conveyed by the system (Prates et al., 2000).

The goal of both methods is to assess the quality of the metacommunication message created by the designer. The difference between the two is that SIM does not involve actual users of the system and their behavior. For this reason, both methods capture different perspectives and aspects of the user-system communication and interaction (de Souza and Leitão, 2009; de Souza et al., 2006). These methods are rooted in the emission and reception of the metacommunication message, similarly to RQ2a and RQ2b.

Salgado et al. (2009) extended SemEng to help designers in communicating with users in a multi-cultural context using conceptual metaphors. Multi-cultural applications include users of different cultures, *i.e.*, users with different values. The proposed metaphors were meant to be used to guide designers when mapping out the interaction of foreign users with systems of different cultures. They symbolize different lenses and perspectives the user can have at interaction time, and designers could use this information accordingly, to make decisions when creating the metacommunication artifact. Although the authors focus on multi-cultural contact instead of user values *per se*, the SemEng perspective of their work is motivating for this work in user values.

The concept of "culture" is too abstract when dealing with values, even though cultures do include values shared by its people. In a later study, Salgado et al. (2011) tested the proposed metaphors by asking designers to re-design an American website with a foreign user in mind. The authors found that the designers gained "awareness of their own cultural biases" during the decision-making processes they went through (Salgado et al., 2011). This awareness avoided the inclusion of perhaps wrong assumptions in the metacommunication message, which could badly affect the user experience. The idea of designers being more careful when dealing with different cultures is encouraging in the context of user values.

User values have not been explored in SemEng as explicitly as in VSD. In this work we plan to delve into the integration of SemEng with user values via MoLIC diagrams (Section 2.4). We aim to evaluate whether this integration results in any difference in the design of the metacommunication message.

2.4 Interaction Modeling

MoLIC is a Modeling Language for Interaction as Conversation, proposed by Barbosa and de Paula (2003) as a SemEng design artifact. It helps designers plan the conversations (*i.e.*, interactions) users may have with the system by generating artifacts that visually show how the user and system interact. MoLIC's main goal is to support the designers in making decisions regarding the system design (Barbosa and de Paula, 2003). The representation of the conversation between the user and the designer through the interface is in the form of a diagram to help structure the information.

To address the lack of consensus on which interaction model should be used when designing a system, Marques et al. (2016) performed a comparative study between MoLIC and CTDM (Comprehensive Tasks and Dialog Model). CTDM represents tasks performed by the user and also the dialog held between

the user and the user interface (López-Jaquero and Montero, 2007). Marques et al. (2016) found that most participants stated a preference for MoLIC over CTDM, and 60% found MoLIC more useful and easier to use for modeling interaction. This further motivates us to use MoLIC diagrams as the artifacts to investigate to answer our research questions.

In Figure 2.1, we have a short example of a MoLIC diagram of a food delivery application. The filled, black circle represents the start of the conversation between user and the designer's deputy. In this case, this happens when the user opens the application. Each rounded rectangle depicts a scene, which represents conversations about specific topics. In each scene, utterances are marked with "u", for user, "d", for designer's deputy, or "d+u", for dialogues between both.

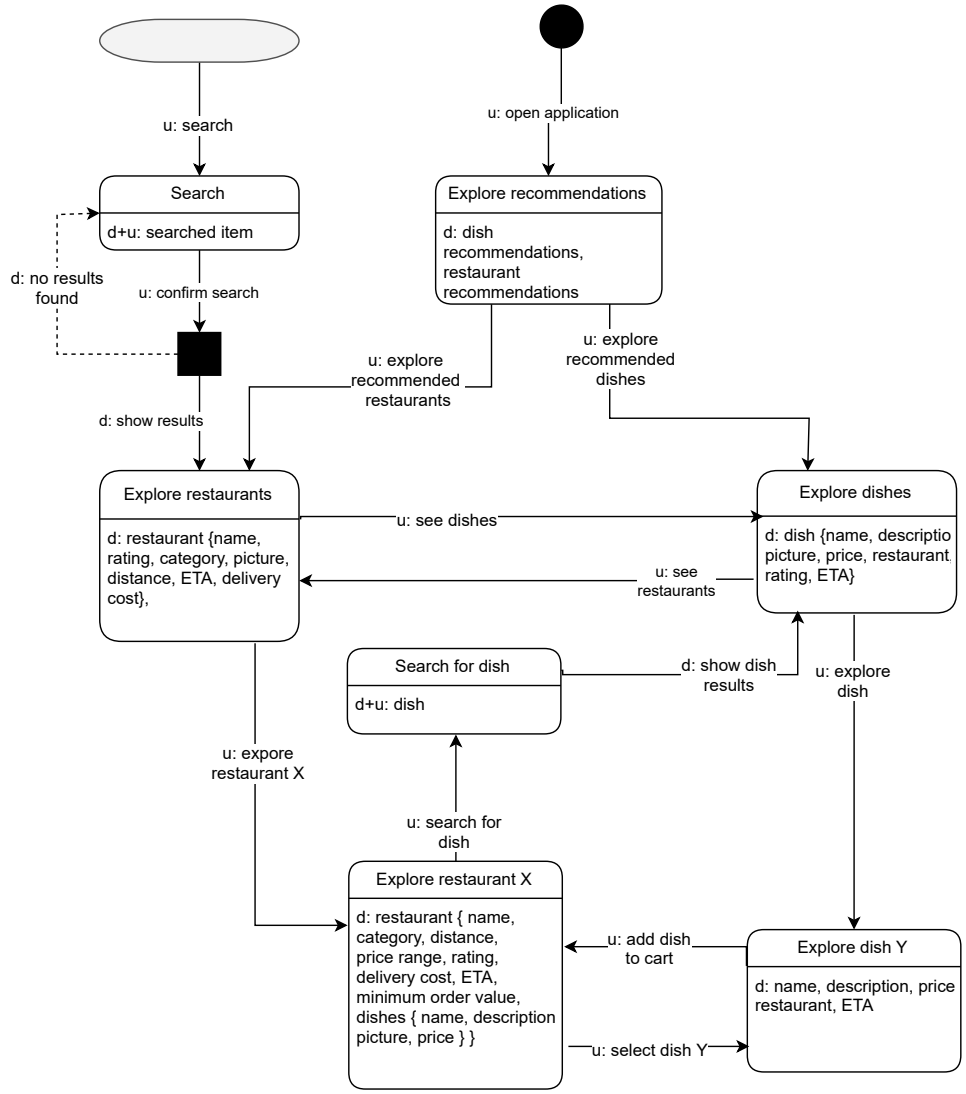

Figure 2.1: Part of a MoLIC diagram of a food delivery application

The scene after the start of the conversation is "Explore recommendations", in which the designer's deputy displays dish and restaurant recommen-

dations. The user, then, can choose between each type of recommendation, leading him/her to a new scene.

It is also possible to search for a specific item. This can be done anytime during the interaction, which is represented by the rounded grey rectangle. This is called, in the MoLIC specifications, "ubiquitous access". Then, after searching for an item, the system performs a certain process, represented by a black box. This means that, after the process is concluded, the designer's deputy needs to decide the next conversation topic. Note, however, that a communicative breakdown may occur, *i.e.*, something may happen that hampers the successful achievement of the goal. The dashed line, which says "d: no results found", represents a *breakdown recovery transition utterance*. In this case, the designer's deputy redirects the user to the search scene, stating that no results were found. If there were indeed results found, the designer's deputy moves on to the next scene, in which he/she shows the user the search results of restaurants.

In the next section, we present the research procedure planned for this work.

3
Research Procedure

As stated in the Introduction, the research questions for this work are: "How can user values be integrated into design rationale?" (RQ1), "How can design artifacts reflect the design rationale taking into account user values?" (RQ2), "Given a design rationale which explicitly references values, do designers embed those values in the design artifacts, explicitly or implicitly?" (RQ2a), and "If a design artifact embeds values, will designers recognize those values?" (RQ2b).

In order to answer these questions, the first step is to evaluate an integration of user values into design rationale (DR). Because there are many approaches to DR, we have selected Questions, Options, and Criteria (QOC) for its simplicity and straightforwardness (MacLean et al., 1991).

In order to evaluate how the integration may affect the creation of design artifacts (RQ2), we will consider MoLIC diagrams as the artifacts to be analyzed. The goal is to see what kinds of differences will appear in the artifacts with and without the integration of user values into the design rationale.

We will[1] conduct two studies with undergraduate and graduate Computer Science students in Rio de Janeiro. The application domain will be a food delivery service, on which the participants will perform design rationale and modeling activities. The first study, which will focus on the impact of user values on DR, is described in Section 3.1 and the second study, which will focus on the impact of user values + DR on design artifacts, is described in Section 3.2. In Figure 3.1, we provide a graphic overview of both studies.

3.1
Study 1: Impact of user values on DR

The first user study focuses on the impact of using user values in a design rationale process. To do that, we will have two groups of participants: the first one will work with user values (Group V) and the second one will not (Group N, as in 'non-value-based', or more 'generic' in comparison). The first study comprises two steps: in the first step, participants will generate a design

[1]Although the studies have already been conducted, we use the future tense in this chapter to reflect the moment when the study was planned but had not yet been carried out.

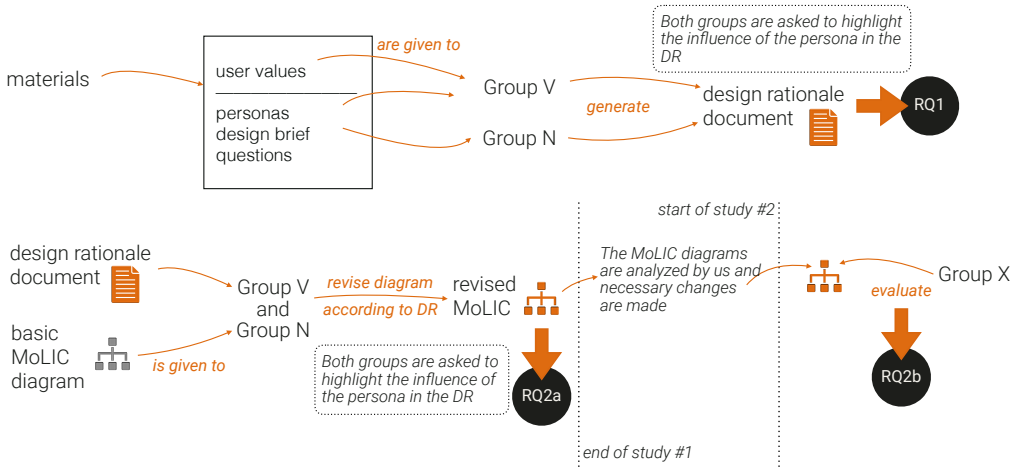

Figure 3.1: Overview of the studies

rationale document. In the second, they will edit a MoLIC diagram to reflect the design rationale they will have documented.

In the following subsections we detail the specifics of this study.

3.1.1 Preparation

Prior to the study, we have created two personas of food delivery applications taking into account user values. One persona has strong values regarding veganism, sustainability, and charity/social work. This persona, named Sofia, has been a vegan for over a year and likes to order food in mobile applications, although she is afraid of the quality of the food and of whether it is actually vegan. She also has strong well-defined environmental concerns: she carries a metal straw and reusable shopping bags with her on the daily, and she recycles her trash and does not like to receive packaging made of foam from restaurants as she cannot recycle it. She goes thrift-shopping and rides her bike to work. She donates monthly to animal protection charities. The value-based persona can be found in Figure 3.2, in Portuguese.

The other persona, Janaína, is more "generic", *i.e.*, was created with no specific values in mind. She often likes to order dinner from her phone, but does not have any dietary restrictions. She is satisfied with her experience on generic food delivery applications. She represents a group of users with functional requirements regarding food delivery. The non-value-based persona can be found in Figure 3.3, in Portuguese.

It is important to note that some participants may be familiar with

Figure 3.2: Sofia - Value-based persona

the concept of food delivery applications. Some may already be users of such services and may have their own values, which can then influence their decisions. To reduce the potential bias of such situation, we will ask participants, at the moment of recruiting, the degree of familiarity with this type of application. We will use this as a criterion to distribute participants across both groups, so that the groups will have similar compositions.

Because of the possible familiarity, we will provide a brief questionnaire to the participants to find their own user values regarding the domain. We believe their values may influence their decisions, so we need to document this possible bias. In this questionnaire, we will also ask their age and their familiarity with food delivery applications. This questionnaire can be found in Appendix A, in Portuguese.

If we get enough participants, we will then define pairs inside each group to get richer results. If not, each participant will perform the activities solo.

We will generate a MoLIC diagram of basic actions a user can perform in a food delivery application, which will be used for the second step of the study. We envision mainly Create, Read, Update and Delete (CRUD) tasks, which will be derived from real food delivery applications.

Figure 3.3: Janaína - Non-value-based persona

3.1.2 Materials

For both Group V and Group N, we will provide a design brief to reveal the software requirements, including a persona, and a list of questions to drive the design discussions and process. As mentioned, for Group N, the persona will be more "generic" than for Group V, as a way to minimize the appearance of user values. Group V's persona, Sofia, carries strong values such as veganism, environmental concerns – recycling, reducing, and reducing waste –, frequent charity donations, eco-friendly transportation, thrift shopping, and, in the context of food delivery applications, reliability regarding food quality. Group N's persona, Janaína, is more bland and generic.

On the second step of the study, we will provide participants with the MoLIC diagram created previously based on existing food delivery applications. It has basic actions related to food delivery, such as "Explore dishes and restaurants" and "Add item to bag". The diagram can be found in Figure 3.4, in Portuguese.

3.1.3 Procedure

We will ask participants from both groups to propose options and criteria for each question, following the QOC method. Since Group V will be given a value-

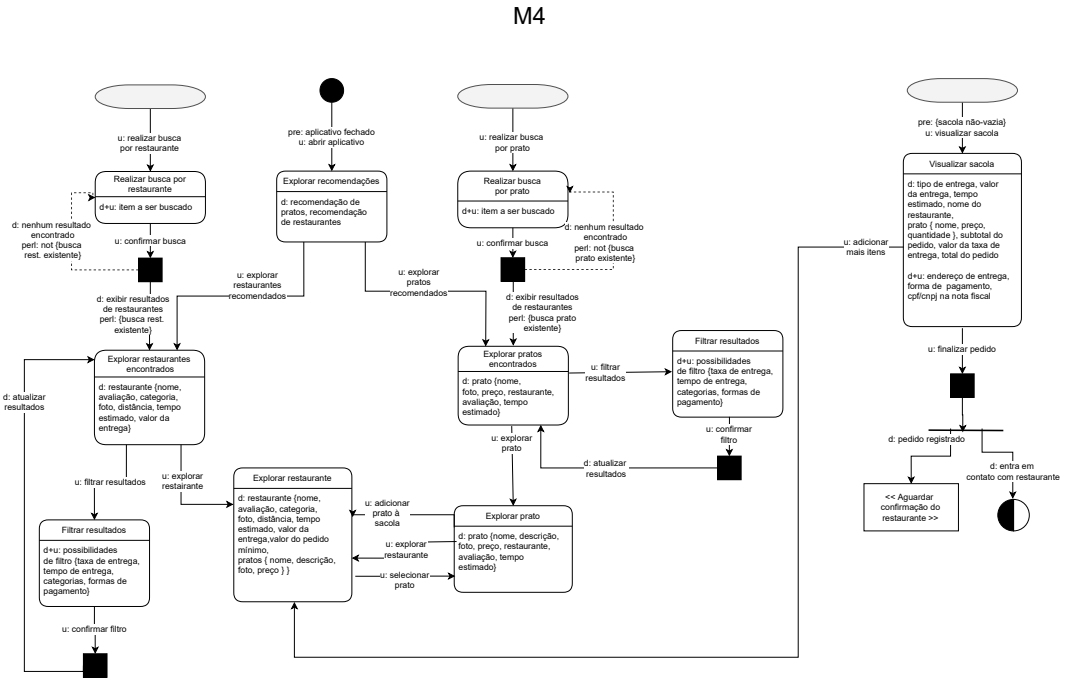

Figure 3.4: Basic MoLIC diagram for a generic food delivery application

based persona, we will ask this group to note whether the persona's own values motivated any options or /criteria. Participants will, then, generate a design rationale document, with their selected options for the questions provided.

We will provide a questionnaire after participants submit the design rationale, in which we are interested in finding whether participants' own values, which were stated before the activities took place, will have impacted their decision-making process. Also, more generally, we want to know how the participants will have chosen from among the options. For Group V, we will also ask how they dealt with conflict between user values, if any (*e.g.*, did they create some kind of hierarchy?).

At the end of this step, participants will have their own design rationale regarding a food delivery service. We will analyze the design rationale documents created by the participants as a way to answer RQ1.

A second step of this study will be the revision and editing of a MoLIC diagram, as a resource to answer RQ2. We will ask participants to alter the MoLIC diagram provided according to the design rationale they generated. For Group V, we will also ask participants to note, in the diagram, when changes were made based solely on the persona's values. We will analyze the revised MoLIC diagrams to answer RQ2a. To answer RQ2b, those diagrams will be evaluated by other participants on Study 2, as discussed in the following section.

In Chapter 4 we go through the participant distribution, as well as the

results from this study.

3.2 Study 2: Impact of DR + VSD on design artifacts

After the first study is concluded, we will perform a second study to assess whether user values made a difference in the MoLIC diagram, *i.e.*, whether they would alter the final software product.

The second study will be conducted with new participants, this time graduate Computer Science students. These participants will be asked to fill out the same questionnaire as the participants in Study 1, which can be found in Appendix A, in Portuguese. Each participant will perform the tasks alone.

3.2.1 Preparation

Prior to this study, we will revise the MoLIC diagrams submitted by the participants from the previous study. We will correct syntactic errors and exclude diagrams that contain too many errors and/or too many misconceptions.

There is a possibility that the MoLIC diagrams turn out too different from one another, which can make it hard to compare. This can happen if participants have clashing ideas regarding the interaction flow, which can be independent of user values. Therefore, before the second study, we will analyze all diagrams and see if any changes and adaptations unrelated to user values need to be made to make them all more comparable. These changes would be documented for reference. We will then choose four variations of the diagrams, settling on two more "generic" (non-value-based) and two value-based.

3.2.2 Materials

Each participant will receive two MoLIC diagrams derived from the first study. One will be from Group V and the other from Group N. As we have four different diagrams, each participant will receive a combination of one value-based and one non-value-based. We want to create different combinations so as to avoid potential bias.

We will also provide the personas created for the first study. Each participant, then, will receive the value-based persona and the non-value-based persona created beforehand for Study 1.

3.2.3 Procedure

Each participant will analyze the MoLIC diagrams they receive. We want answers to the following question: *Do the user values embedded in the MoLIC diagrams make any noticeable difference in the design process?* Thus, we will first ask the participants to describe the potential users for the systems described in the diagrams they will have received. After their description, we will introduce the concept of *user values* and ask them to list what values the users of the systems may have regarding food delivery applications. These first two questions will be answered based solely on their perception of the diagrams, with no specific prompts regarding specific values.

Next, because we will have provided a persona in the beginning of the first study, we are curious in finding to which degree each MoLIC diagram is aligned with each persona. To assess this, we will ask participants to associate the personas with the diagrams and to provide a scale of alignment between both artifacts, to assess to what extent a MoLIC diagram is perceived as suitable for each persona. The point of this assessment is to see whether the resulting artifact will be perceived as appropriate for its respective persona. The value-based persona contains specific information regarding sustainability, veganism, and charity work. Some of these values will be reflected in the value-based MoLIC diagrams.

We will also ask participants to associate the MoLIC diagrams with user values, from the list in the questionnaire provided before the first study. We want to see whether they can identify what values are present in each MoLIC diagram. This will be done via a questionnaire, which will also feature the alignment scale between each MoLIC diagram and each persona. This questionnaire can be found in Appendix D, in Portuguese.

Finally, we will provide a profiling questionnaire – the same used for the first study – and use the results of Study 2 to answer RQ2b. We want to find whether participants will have identified the user values that were embedded in the value-based MoLIC diagrams.

In Chapter 5, we discuss the MoLIC diagrams used, the participants' profile, and report the results of this study.

4
Study 1

This chapter describes the results of Study 1. We first go over the participants' profiles and distribution in Section 4.1, then move on to report their responses in Section 4.2. Finally, in Section 4.3, we discuss the results and answer Research Questions 1 and 2a. This analysis was done based only on the collected data; we were not able to check our inferences with participants.

4.1
Participants' groups

The first study was performed with 15 undergraduate Computer Science students during a Human-Computer Interaction class, in Rio de Janeiro. There were 13 Brazilian students and two French exchange students. We explained the context of the study to the students, asking their permission to use the data for research purposes. The students then signed an informed consent form, which was not shared with their professor for privacy and ethics reasons. We then provided a brief questionnaire in which we asked their age, how often they use food delivery applications, and we asked students to assess, in a 7-point Likert scale, how important certain issues are when ordering food in mobile applications. The issues were: Price, Familiarity with the restaurant, Hygiene, Type of packaging used by the restaurant, Estimated time of delivery, Type of food, and Data security. We considered these aspects as common values regarding food delivery applications, and some of them were included in the value-based persona. The reason we asked students to rate the importance of these aspects was because their own personal bias could influence the way they performed the study activities. Documenting this could be beneficial when analyzing the study results. This questionnaire can be found in Appendix A, in Portuguese.

After the students filled out the forms, we separated them into pairs, to balance the degree of familiarity with food delivery apps and the students' values. For instance, students who had never ordered food online were paired with students who often do. However, the majority of students claimed to order food frequently, so some pairs featured students with similar habits. In these cases, we looked into their values and tried to choose pairs with different

value preferences. The pairs could perform the activities in either Portuguese or English.

After grouping the students, we had seven pairs and one single student. We randomly assigned four pairs to Group V, *i.e.*, the group with the value-based persona, and three pairs and a single student to Group N, *i.e.*, the group with the more "generic" persona.

The full profile of the participants can be found in Table 4.1 and Table 4.2. The Pair column in Table 4.1 indicates the resulting pair code, V representing Group V and N representing Group N. N4 was the participant who worked individually, belonging to Group N. In Table 4.2, the scores were given by participants in a Likert scale, 0 (– – –) being "I do not care about this aspect" and 6 (+++) being "I care deeply about this aspect". Note that P2 did not fill out the scale, claiming they never use food delivery applications. Other participants who also claimed to never using these types of services – P3, P8, and P10 – did, however, fill out the scale.

Table 4.1: Participant profile

	Pair	Age	Frequency of food delivery app usage
P1	V1	< 20	Every week
P2	V1	21-30	Never
P3	V2	21-30	Never
P4	V2	21-30	Every week
P5	V3	< 20	Every month
P6	V3	< 20	Every trimester
P7	V4	< 20	Every month
P8	V4	< 20	Never
P9	N1	< 20	Every month
P10	N1	21-30	Never
P11	N2	21-30	Every month
P12	N2	31-40	Every month
P13	N3	21-30	Every month
P14	N3	< 20	Every week
P15	N4	21-30	Every week

4.2 Participants' responses

In this section, we describe the participants' performance in the assigned activities.

In the first activity, the participants had to create a design rationale document using the following questions:

– Q1. How can the user find what to order in the application?

- Q2. What kinds of search should the application allow? That is, for what can the user search?
- Q3. What types of filters should the application allow for the search results?
- Q4. What kinds of recommendations should the application give the user? Based on what?
- Q5. What information about each restaurant should be given to the user?
- Q6. What information about each dish should be given to the user?
- Q7. What information about each order should be given to the user?

The questions should be answered according to the persona each pair received. We asked participants to note in the design rationale document when the persona specifically motivated the decision-making process.

Since the students had not previously had design rationale lessons, we provided a short explanation of the basic structure of the Questions, Options, and Criteria (QOC) method. For each question provided, the students should provide options and criteria, according to the persona received. We emphasized the importance of the persona and that the food application they were designing should be targeted for that specific persona.

We mentioned in Section 3.1.3 that we would provide a questionnaire after this step to understand their motivations when answering the questions, *i.e.*, whether their own values may have influenced their decision-making process. Unfortunately, however, due to time constraints, we did not perform this step. Some participants took longer than expected to perform the tasks

Table 4.2: Likert scale of aspects that may influence participants' activities regarding ordering food in mobile applications

	Price	Familiarity	Hygiene	Packaging	Time	Food type	Data Security
P1	+++	–/+	– – –	– – –	–	+	– –
P2	NA	NA	NA	NA	NA	NA	NA
P3	+++	–/+	++	– –	+	+++	–/+
P4	+++	++	+++	+++	+++	– – –	+++
P5	+++	–/+	++	–/+	–	– – –	+++
P6	+	++	+++	–	+++	++	+
P7	+++	+	+++	–/+	+	– – –	++
P8	+	–/+	++	–/+	++	+++	++
P9	++	+	–/+	– –	++	+	–/+
P10	+++	+	+	–	++	+	+
P11	++	– – –	+++	+	+	–	++
P12	+++	++	+++	+++	+	+++	+++
P13	++	+	+++	+	–/+	+	+++
P14	++	–	– –	– –	–	++	– –
P15	+++	– – –	+++	– – –	–/+	–/+	++

because they did not fully understand the activities. We could have sent written instructions prior to the study, to make sure participants were aware of the tasks. The questionnaire could have been used to perform deeper analyses of the participants' performance and rationale, which is a limitation of this study.

The second activity was the revision of a MoLIC diagram of a food delivery application to reflect the decisions made during the design rationale process. In this revision, participants should alter the interaction flow as they saw. They could also change the signs exchanged between the user and the designer's deputy (*i.e.*, the user interface). The students had had MoLIC and SemEng lessons prior to the study, but we also provided a short document with a quick reference to the MoLIC notation.

4.2.1 Answers to Questions Q1 to Q7

In this section, we summarize the answers participants gave to questions Q1 to Q7, presented in Tables 4.3 to 4.9. In the tables, each row represents an item listed as an option recorded by at least one participant in the QOC rationale, and each column represents a pair of participants. Note that we did not provide a list of possible options. In the cells, X indicates that the pair listed the item as an option, but did not select it as an answer; X in boldface indicates that the pair listed the item and selected it as an answer; and X* indicates that the pair listed the item, selected it as an answer, and associated it with the given persona. Pair names in gray indicate that the pair failed to provide a valid QOC rationale for the question. For each table, we noted what options we believe are related to the values present in the value-based persona (*e.g.* sustainability).

Table 4.3: Participants' answers for Q1 (How can the user find what to order in the application?).
[1] can be related to the veganism value in the persona (dietary restrictions).
[2] related to the trustworthiness value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Search[1]	**X***	**X**	X	**X**	X		**X**	X
Categories[1]	X		**X***	X	**X***			
Recommendations	X			X	X		X	
Previous orders[2]	X				X		**X**	
Filters[1]		**X***		**X**				
Daily offers			X					X
Most popular		X						
Highly rated		X						

In Q1 (*How can the user find what to order in the application?*), the most

popular options were search, categories and recommendations. Although they may seem like generic solutions to Q1, search and categories do apply to the value-based persona's needs. Via search and category exploration, a vegan user can search for famous vegan dishes or explore the vegan category. Filters were also mentioned by two pairs in Group V, but only one related it to the persona. We see "filters" as a loose term for a feature that could help the value-based persona find what to order in the application, considering the possibility of having filters targeted to their needs.

In Q2 (*What kinds of search should the application allow? That is, for what can the user search?*), the only option related to the value-based persona was "food type", since Sofia is vegan, and it was chosen by two pairs in Group V and two pairs in Group N. The first two, however, associated the option with the persona, whereas the last two, in Group N, did not. "Restaurant" was the most popular option, having no relation with any persona. We also consider "dish" as an option related to the value-based persona, since a vegan user could search for a specific vegan dish, but Group V did not create this link.

Table 4.4: Participants' answers for Q2 (What kinds of search should the application allow? That is, for what can the user search?).
[1] can be related to the veganism value in the persona (dietary restrictions).

	V1	V2	V3	V4	N1	N2	N3	N4
Restaurant		**X**	X	**X**	X		X	
Food type[1]		**X***	**X***				X	X
Dish[1]		**X**		X	**X**		**X**	
Discounts			X				X	
Price range								X
Delivery time								X

In Q3 (*What types of filters should the application allow for the search results?*), "delivery time" and "price" were popular options, having no obvious connection to either persona. We believe this bias was due to the participants' previous knowledge and usage of food delivery applications. All pairs in Group V, though, associated some options with their persona, whereas Group N did not. For Group V, there were "rating", which is related to trustworthiness, "category" and "dietary restrictions", related to veganism, and "packaging type", related to sustainability. Out of these options, only "category" also appeared as an option for a pair in Group N, but was not chosen as an answer.

In Q4 (*What kinds of recommendations should the application give the user? Based on what?*), the most popular options were "highly rated" and "discounts", neither of which directly related to the personas. The pair V2,

Table 4.5: Participants' answers for Q3 (What types of filters should the application allow for the search results?).
[1] related to the trustworthiness value in the persona.
[2] related to the veganism value in the persona.
[3] related to the sustainability value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Delivery time	X	**X**	X	X	**X**		X	X
Rating[1]	**X***		**X***	**X**	X		X	X
Price	X	**X**	X		X		**X**	X
Category[2]	**X***	**X***						X
Dietary restrictions[2]	**X***							
Packaging type[3]				**X***				
Most popular		X						
Distance					X			
Delivery fee							X	
Payment options							X	
Discounts								X
Favorite restaurants								X

however, associated "highly rated" to the value-based persona, but it is unclear why. "Previous orders" and "packaging type" were options related to the persona, and two pairs in Group V made that connection.

Table 4.6: Participants' answers for Q4 (What kinds of recommendations should the application give the user? Based on what?).
[1] related to the trustworthiness value in the persona.
[2] related to the sustainability value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Highly rated		**X***		**X**	X			X
Discounts		X			X		**X**	X
Previous orders[1]		**X***	**X***		**X**			
Delivery time			X	X				
Most popular		X					X	
Packaging type[2]				**X***				
Average price of user's orders			X					
New in					X			
Personalized choices							X	
Favorite dishes								X
Favorite restaurants								X

In Q5 (*What information about each restaurant should be given to the user?*), there were a few options that Group V linked with their persona: "food type", "sustainable delivery", "eco-friendly", "vegan/vegetarian", and "packaging type". We also consider "rating" as a good option for the trustworthiness value present in the persona, but Group V did not make this connection. N1,

from Group N, associated "food type" with the non-value-based persona, which is unclear.

Table 4.7: Participants' answers for Q5 (What information about each restaurant should be given to the user?).
[1] related to the trustworthiness value in the persona.
[2] related to the veganism value in the persona.
[3] related to the sustainability value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Rating[1]		X	X		X		**X**	
Food type[2]		**X***	X		**X***			
Price range		X	X				**X**	
Delivery time			X		X		**X**	
Distance		**X**			X			
Sustainable delivery[3]		**X***						
Eco-friendly[3]		**X***						
Vegan/vegetarian[2]		**X***						
Packaging type[3]			**X***					
Delivery fee							**X**	
Payment options							**X**	
Working hours							**X**	
Available dishes					X			

The participants gave many options to Q6 (*What information about each dish should be given to the user?*), most of which were provided by only one pair. Some options were associated with Group V's persona, but this association is not clear (*e.g.*, "allergy alert" and "information on salt/sugar"). N1 also associated "calories" to their non-value-based persona, which was unexpected.

Finally, for Q7 (*What information about each order should be given to the user?*), there were only two options associated with the value-based persona, which were "dietary restrictions alert" and "transportation type", the latter being related to sustainability (for instance, bicycle delivery). Pairs in Group N did not associate any options with their persona.

We can note that most of the answers associated to the personas lie in the pairs in Group V, indicated in the tables with an asterisk. Some values appear in both groups, such as concerns with pricing and delivery time, which may indicate participants' biases. These features are common for food delivery applications, so we believe participants tended to incorporate their previous knowledge and application usage into their design rationale documents. However, despite this issue, Group V presented richer results in terms of modelling an application targeted to a specific user, showing concern with veganism and sustainability. Group N tended to remain superficial and

Table 4.8: Participants' answers for Q6 (What information about each dish should be given to the user?).
[1] related to the veganism value in the persona (dietary restrictions).
[2] related to the trustworthiness value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Description	X		**X***		**X**		**X**	
Price	X		X		**X**		**X**	
Restaurant	**X***						**X**	
Name	X						**X**	
Image	X		X					
Weight		X			X			
Dietary restrictions alert[1]	**X***							
Rating[2]	**X***							
Allergy alert		**X**						
Vegan/vegetarian[1]		**X***						
Info on sugar/salt		**X***						
Dish prep time	X							
Types of meat[1]		X						
Food type[1]			X					
Calories					**X***			
Nutritional info							**X**	
Ingredients[1]					**X**			
Customization options[1]							**X**	
Delivery time							**X**	
Favorited					X			

very rarely selected options due to any association to the non-value-based persona. Only the pair N1 explicitly chose options based on their persona, whereas all pairs from Group V did so at some point in the study.

4.2.2 MoLIC diagram changes

After creating the design rationale, participants were asked to alter a MoLIC diagram to reflect their decisions. In Table 4.10, we summarize the changes made by both groups and the full diagrams of Group N and Group V can be found in Appendix B. No pair explicitly noted that the changes made were due to their respective persona. However, we can infer that some changes were only made because of the values present in the value-based persona, such as the inclusion of new signs of packaging type used by the restaurant and of dietary restrictions concerns.

We also find that Group N tended to add more generic signs and utterances, such as price information and ratings. Group V tended to stick to the needs of their persona, without adding many more extra signs and

Table 4.9: Participants' answers for Q7 (What information about each order should be given to the user?).
[1] related to the veganism value in the persona (dietary restrictions).
[2] related to the sustainability value in the persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Order summary	X		X		**X**		X	
Order price	X	**X**			**X**		X	
Delivery time		**X**	X				X	
Delivery fees	X	**X**					X	
Payment options	X				X		X	
Restaurant	X				**X**			
Delivery/order status					**X**		X	
Dietary restrictions alert[1]	**X***							
Transportation type[2]		**X***						
Delivery options	X							
Tip	X							
Date/time					X			
Restaurant contact info					X			
Coupon							X	
Delivery address							X	

features.

Unfortunately, N2's MoLIC diagram had no sufficient modifications to be featured and analyzed in this step of the study.

Since the alteration of the MoLIC diagram was the last step of Study 1, it is possible that participants were tired or had little time to perform this activity. In a different scenario, perhaps the results would have been richer.

4.2.3 Group V Results

As mentioned in Section 3.1.2, Group V's persona had values of veganism and overall environmental concerns. We expected the design rationale to reflect the persona's characteristics and values.

The pair V1 had one of the most interesting results among Group V. They included a notion of dietary restrictions among their options to a few questions, such as Q3. This idea generated a "restrictions alert", which would inform the user if any of their dietary restrictions was being violated within their order. They noted that the user should inform the application, before placing an order, of their dietary preferences, which the system would then use to provide the possible alert. V1 also included filters by category and restaurant rating, using the same criterion of dietary preferences. We believe the rating option was selected because the persona claimed not to trust most

Table 4.10: Changes in the MoLIC diagram made by participants.
[1] related to the value-based persona.

	V1	V2	V3	V4	N1	N2	N3	N4
Generic signs	X		X		X		X	X
Generic transition utterances	X				X		X	X
Packaging signs[1]	X	X	X	X				
New scene: Search by category[1]	X			X	X			
New scene: Explore discounts					X		X	
New scene: Fill out dietary restrictions[1]	X							
New scene: See popular orders							X	
New scene: Add observation to existing order							X	
Vegan/vegetarian signs[1]		X						
Restaurant recycling signs[1]		X						

restaurants (trustworthiness value). Also, when answering Q5, they included an option of informing the user whether the restaurant was "environmentally responsible", a characteristic they did not specify further. Unfortunately, V1 only answered 4 out of 7 questions.

In the MoLIC diagram, V1 included an option to search by category. They also allowed the user to explore recommendations based on previous orders and favorite dishes from the app development team. They did not specify the reasoning for these additions.

The pair V2 considered search and filter by food type (vegan) for the persona. They also decided that the application should provide recommendations based on highly rated and most popular restaurants, which was not related specifically to the persona. We found this pair selected options based on their own reasoning more than on the persona's values. For instance, on Q3, relative to types of filter, besides choosing "food type" they also chose price and delivery time/distance, using a criterion of "matches users' habits". The persona did not contain any information concerning their attention to pricing and delivery time. However, for Q5, V2 came up with an interesting solution. They decided that the application should inform the user on whether the restaurant provides a "sustainable delivery", on whether it is vegan/vegetarian, on whether it is "eco-friendly", the type of food the restaurant provides and also the distance. The criteria related to the persona were: "helps the environment" and "matches user's beliefs". For Q6, V2 decided that, for each dish, the application should have an allergy alert, "to protect the user", while also listing every information regarding sugar/salt and whether it is vegan/vegetarian. They selected the sugar/salt option as being based on the persona, but the reasoning behind this is not clear.

In V2's MoLIC diagram, they altered only the signs in the dialogues exchanged between user and designer's deputy, *e.g.*, type of packaging used by

the restaurant, whether it is vegan, and whether the restaurant recycles. The interaction flow remained the same.

V3 decided that the best way to help the user find what to order (Q1) is through food categories, as they provide "better precision". For the search engine (Q2), they chose search by food type, and for the filters (Q3) they picked only highly rated restaurants to improve trust, which is valued by the persona. This pair did not consider veganism or vegetarianism when performing the activities, but included packaging type as information to be displayed about each restaurant (Q5). They opted for providing recommendations based on previous orders (Q4), also using trust as a criterion. They only selected one option per question, although we reminded the participants that they could select multiple. Questions 6 and 7 had more generic options and criteria. Like V2, V3's MoLIC did not go through many changes, namely including and removing signs in scenes.

V4's design rationale was incomplete. The relevant content targeted for the persona included highly rated restaurants, type of packaging filter and search by category (such as vegan). Their MoLIC diagram, similarly to V1, included a new search by category. They listed as potential categories: pasta, vegan, vegetarian, and pizza. They also added a packaging type filter on the existing "Filter results" scene.

4.2.4 Group N Results

For Group N, the persona was more bland and generic. It described a woman who lives by herself, has a regular nine to five job and for dinner likes to order food online. She goes to the gym and likes music. On the weekends, she likes to watch comedy shows and hang out with friends and family.

We tried to create a persona with common likes and habits, who orders food frequently. She does not have strong convictions like Group V's persona, who strongly believes in veganism and in protecting the environment through different actions. We may say that Group N's persona is *bland*.

The pair N1 had a curious perception of their persona. For Q1, they chose to show the user different options through categories, and the criterion for this option was "diet". Perhaps because their persona goes to the gym they inferred that she is on a diet and would like to order food accordingly. They had other options for Q1, such as recommendations, order history, and search, but they chose the one they believed was most adequate for their persona because of her "diet". This was unexpected, as we did not mean to include any values for this persona. For Q6, N1 also used the "diet" criterion to choose an option. They

decided that each dish should have its calories listed, so that people on a diet can benefit from that. These were the only instances in which N1 specified a reasoning based on the persona's characteristics for choosing the options. The other questions were apparently answered based on N1's own preferences and previous experience.

The edits N1 made to the MoLIC diagram included a change in the start of the conversation with the system, from the scene "Explore recommendations" to a new scene "See categories". This makes sense, as they chose categories as the answer to Q1. They also included a "See discounts" scene, despite not having selected any options regarding discounts on their design rationale. Also, as mentioned before, they included calories in the "Explore dish" scene.

N2 had the most incomplete design rationale, with only two questions. The first question had four options and the second had three, and no option had any criteria. They did, however, make a few changes in the MoLIC diagram. They combined the two different types of search (by dish and by restaurant) into one single type, while also including filter options on the new search scene. The search scene, then, returns results for both restaurants and dish. This change does not reflect their design rationale and does not seem to provide great impact on the overall flow.

N3 did not highlight any options on their design rationale as being motivated by the persona. Overall, they prioritized pricing and discounts. This may have happened as P13 and P14 claimed to care a lot about prices when ordering food online, as seen in Table 4.2, so their personal bias may have influenced their decision-making process. They also answered Q6 with basic options, like restaurant name, but decided to include nutritional information, which is not included in most food delivery applications. In general, though, their selected options for the questions provided seemed based on common sense, not on the persona itself.

In the MoLIC diagram, N3 included three new scenes: "Explore discounts", "Explore popular orders", and "Add observation to order". Similar versions of these scenes can be found in common food delivery applications today. Both participants of N3 claim to use these types of applications often (monthly and weekly), as seen in Table 4.1.

The final participant, N4, performed the activities alone. Their design rationale was incomplete, featuring the first four questions, but the participant failed to select the options. Therefore, it is difficult to interpret the design rationale, leaving only the MoLIC diagram as a document for us to analyze. N4 created a possibility to save restaurants as favorite and created a scene

for exploring discounts. They also added a "popularity" attribute to both restaurants and dishes.

4.3 Discussion - Study 1

We see a clear difference between the results between Group V and Group N. The first presented more specific features targeted to vegan/vegetarian and environmentally-conscious/sustainable users. Packaging type seemed to be a valid concern for most pairs in Group V, and many also considered it important to split dishes and restaurant into categories, to help vegan users.

Group N seemed to base their decisions more on personal experience and usage of food delivery applications. Discounts and pricing appeared as common themes during the design rationale, often being selected as must-have options. In Table 4.2, it is clear that Price is all-around the main factor that influences the participants' behavior in ordering food online. Most (3/4) of Group N added discount-related features into both their design rationale and MoLIC diagram, whereas no pair in Group V did. Neither Group V nor Group N's persona mentioned concerns about saving money, but Group N still included this feature. We believe this happened because of the lack of character traits in Group N's persona, which led them to apply their own presumptions to the design activities. Group V, however, appeared to be more focused on the values of their persona and less on their own personal values.

4.3.1 Answering RQ1

The first Research Question is "How can user values be integrated into design rationale?". This can be answered by analyzing the design rationale documents generated by the participants of this first study.

We asked participants to highlight when the persona's values were taken into account as they came up with options and criteria for the questions provided. They did that by underlining, circling or by writing "persona" or simply "p" near the corresponding option/criteria. All pairs in Group N did a variation of this, along with only N1 of Group N. As explained in Section 4.2.4, N1 used "diet" as a criterion based on their perception of the non-value-based persona.

An example of a QOC instance by the pair V2, for Q5, can be found in Figure 4.1. Three options were motivated by the persona (Type of food, Sustainable delivery, and Vegan/vegetarian). They also considered the option

"Eco-friendly", which we believe was also motivated by the persona, despite not being explicitly flagged as such.

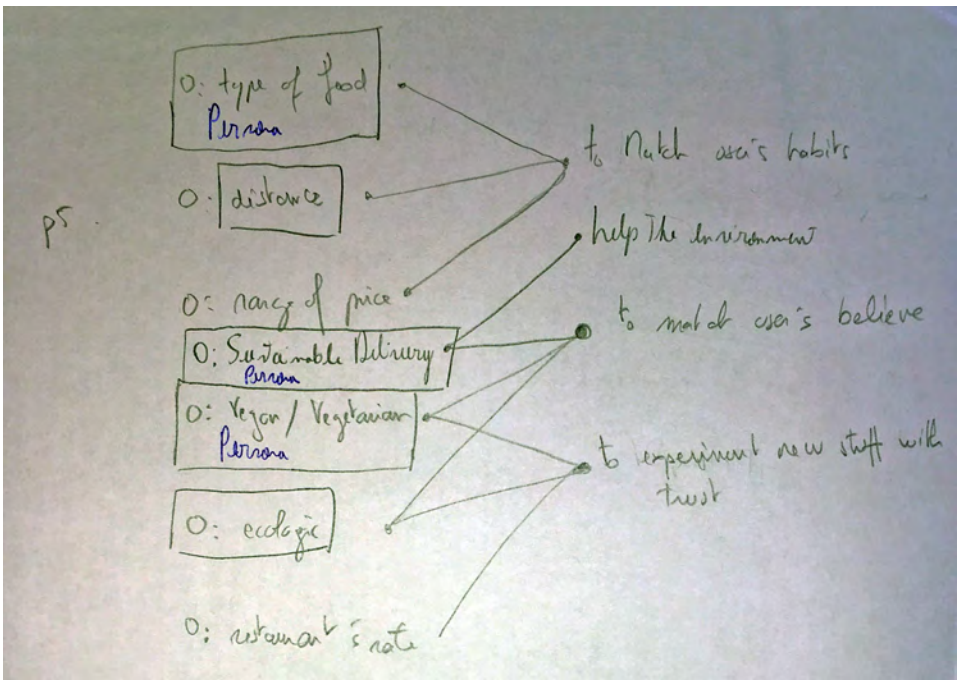

Figure 4.1: QOC instance by V2, highlighting the use of the persona's characteristics and values

An example of a QOC instance for Q5 by N3, in Portuguese, can be found in Figure 4.2. They did not highlight any option or criteria as being motivated by their persona. This was somewhat expected, as the persona handed out to group N was more generic than Group V's. This behavior, apart from N1, appeared to be recurrent among Group N, listing options and criteria seemingly based on their own perceptions of food delivery applications. For instance, the options selected in Figure 4.2 were "Restaurant rating, Delivery time, Payment options, Restaurant hours, Delivery fee, and Average price". We consider these options as being basic and independent of user values, differing from the options considered by V2 in Figure 4.1, which were deeply value-related.

This contrast between the two groups shows us that value-laden personas tend to drive a more personalized approach by designers in decision-making processes. The design rationale documents created by Group V had more specific options and criteria for the questions provided, whereas Group N mostly provided generic options and criteria. The integration of user values into design rationale was successfully achieved by Group V, with different pairs using different notations for highlighting the usage of the persona's characteristics and preferences. By asking the participants to highlight, in some way, this usage, we believe the participants were able to perceive the importance of the persona's wants and needs. That is, options and criteria

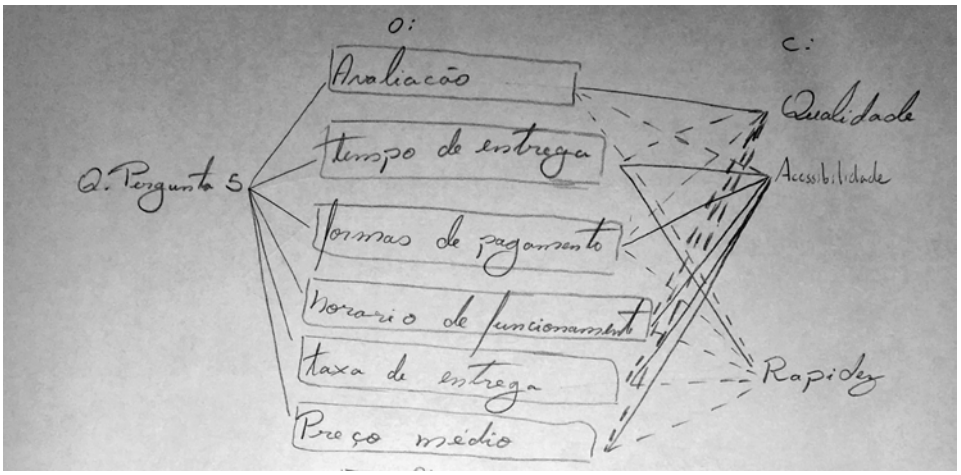

Figure 4.2: QOC instance by N3, with no highlights of the use of the persona's characteristics and values

that were based on the persona tended to be chosen and used more frequently than those which were not. This generates a software design more targeted to the end users and their values.

We found that the notation for highlighting when the persona influenced their decisions made little to no difference between groups. Their own way of emphasizing this influence, whether by underlining, circling or, as seen in Figure 4.1, by writing "Persona" under the options, made the participants, *i.e.*, the designers, aware of their users' needs and values. Therefore, as an answer to RQ1, the integration of user values into design rationale can be made with various notations, as long as the designers understand what is being conveyed and how the users' values affect their decision-making process during design rationale documentation. Future work can include the formalization of a technique for incorporating user values into design rationale processes.

4.3.2 Answering RQ2a

The second Research Question is "How can design artifacts reflect the design rationale taking into account user values?", and was unfolded into two subquestions. The first subquestion is "Given a design rationale which explicitly references values, do designers embed those values in the design artifacts, explicitly or implicitly?". It can be answered by analyzing the MoLIC diagrams altered by the participants, as seen in Table 4.10.

We found that there were no *explicit* values embedded into the MoLIC diagrams, as participants did not make any notes on whether their persona motivated the changes or not. Nonetheless, we found that the changes made by Group V were more appropriate for their persona than the changes made

by Group N, which seemed to be based on the participants' own perceptions, showing a bias. The communicability and emission of user values for Group V, thus, was subtle, and Study 2 will look into the reception of this subtlety.

Since most of Group N's design rationale documents did not explicitly reference values (apart from N1 in a few questions), it is fair that their MoLIC diagram also did not reference any values. Group V, though, had design rationale documents with references to values present in their persona. Some of these references did persist on the MoLIC diagrams, like the packaging signs included by all pairs in Group V. This persistence shows us that at least some values referenced in design rationale documents are indeed embedded in design artifacts by designers, in an implicit way. The subtle addition of packaging, vegan, and recycling signs occurred due to the design rationale taking into account values, and made the overall food delivery application more appropriate for its user. Similarly to the QOC, future work includes an update on the MoLIC notation to formally incorporate user values.

When designing a solution for a specific user with specific values, it seems promising to perform design rationale activities with these values in mind. As seen in this study, at least some values persist in design artifacts, which did not happen for a non-value-based user as seen in Group N.

Study 2 will help us assess with more clarity whether the inclusion of user values in design artifacts are actually recognized by other designers.

4.3.3 Preparation for Study 2

For Study 2, we are interested in assessing the impact of design rationale with user values on design artifacts, more specifically to know whether designers recognize user values in design artifacts that do embed values, *i.e.*, whether the reception of user values and their communicability is successful. To do that, we revised the MoLIC diagrams that resulted from this study in order to settle on four different diagrams: two non-value-based and two value-based.

In Chapter 5 we discuss the results of Study 2.

5
Study 2

In this chapter, we analyze the results of Study 2. In Section 5.1, we discuss the materials used in the study, which were derived from Study 1. Section 5.2 describes the volunteers who participated in Study 2. In Sections 5.3, 5.4, 5.5, and 5.6, we discuss in depth the qualitative remarks made by participants regarding the MoLIC diagrams M1, M2, M3 and M4, respectively. This analysis was done based only on raw data, we were not able to check our inferences with participants. In Section 5.7, we analyze the quantitative results from Study 2, derived from the questionnaires that were provided. Finally, in Section 5.8, we discuss the overall results from Study 2 and answer Research Question 2b.

5.1
Preparation and Materials

Based on the changes made on the MoLIC diagrams during Study 1, we created four new diagrams: two value-based (M1 and M2) and two non-value-based (M3 and M4). These diagrams can be found in Appendix C. Each participant received one diagram from each group.

M1 is the most different out of the four, as it contains a whole new interaction flow of recording dietary restrictions, with new scenes. This idea came from the pair V1 from Study 1. This dietary restriction model was inspired by Sofia, the value-based, vegan persona. It also displays, for each restaurant, the type of packaging used, whether the delivery mode of transportation is "sustainable" (*e.g.* bike), and a bio section. For each dish, there is a Boolean piece of information that states whether it is vegan and/or vegetarian.

M2 does not contain new scenes, only new signs in the dialogues between designer and user. For restaurants, there is a Boolean information regarding whether that restaurant supports any charity. There is also a filter by "category (*e.g.* gluten free and vegan)". For each dish, the user has customization options. In the "shopping cart" section of the application, the user can also customize what packaging is to be received and can choose to donate to the charity supported by the restaurant.

M3 and M4 are value-neutral. The only difference between them is that

M3 has a new scene, "Explore discounts". This was inspired by the pair N3 from Study 1, who was in the "non-value-based" group. Since we noticed that many groups from Group N considered pricing as an important factor when ordering food, we decided to add this scene. Other than that, it is the same as M4, which contains generic interaction and information necessary to order food, as seen on most food delivery applications today.

We performed two pilots prior to the study with graduate students to define its structure, composed of the following tasks:

T1: Participants were asked to describe the potential users of each MoLIC diagram.

T2: After providing a definition of user values, participants were asked to list the user values reflected on each MoLIC diagram.

T3: After providing the personas, participants were asked to align each persona with each MoLIC diagram, through a questionnaire.

T4: Participants were asked to grade how strongly each MoLIC diagram complies with each factor regarding food delivery applications, through a questionnaire.

The factors from T4 were: Price, Familiarity with the restaurant, Hygiene, Type of packaging used by the restaurant, Estimated time of delivery, Type of food, and Data security. These factors were the same from the profiling questionnaire used in Study 1, as seen in Chapter 4, which was also used at the end of Study 2. The questionnaire used for T3 and T4 can be found in Appendix D, in Portuguese.

After T4, we gave participants a brief questionnaire about their own experience regarding food delivery applications. This questionnaire is the same used in Study 1 and can be found in Appendix A, in Portuguese. We have used this questionnaire to assess possible biases in the participants' answers.

5.2 Participants

We conducted the activities with 14 new participants, *i.e.*, none had participated in Study 1. Each participant performed the activities individually, in a small room with one researcher. They signed an informed consent form in which they authorized us to record the audio to the conversation.

Their age ranged from 21 to 40 and they were frequent users of food delivery applications – either monthly or weekly. All participants are in the Computer Science field, and most are in graduate school. One is an undergraduate and one is an MSc currently not in school.

Some participants had prior experience with MoLIC diagrams, but most had none. Since they only had to read the diagram, not modify it, we provided a brief explanation of the notation before the study and explained that participants were free to ask any questions regarding the notation throughout the experiment.

Each participant received a pair of MoLIC diagrams, either: M1 and M3, M1 and M4, M2 and M3, or M2 and M4, as seen in Table 5.1.

Table 5.1: Participants and MoLIC diagrams received

	M1	M2	M3	M4
P1	X		X	
P2	X			X
P3		X	X	
P4		X		X
P5	X		X	
P6	X			X
P7		X	X	
P8		X		X
P9	X		X	
P10		X		X
P11		X	X	
P12		X		X
P13	X		X	
P14	X			X

We analyze the qualitative results regarding each MoLIC diagram separately in the following sections, broken down by task. These results originate mainly from the participants' verbal remarks during the activities, and are related to the identification of user values in the diagrams. Next, we analyze the quantitative results, obtained from T3 and T4.

5.3 M1 - Qualitative results

For M1 (the one with the most noticeable user values), we have participants P1, P2, P5, P6, P9, P13, and P14. The full table of remarks made by participants, broken down by task, can be found in Appendix E.

5.3.1 M1 - T1

For the first task, participants were asked to describe the potential users of the system modelled by M1. Some participants described this user with respect to circumstantial issues – for example, P2 described the user as someone with

no time to cook, based on his own experience when ordering food. P14 also self-identified with the potential users of M1, claiming it to be "just like any other food delivery app". P6 mentioned the ratings as a helpful way for the user to know more about the dish he/she is ordering.

Five out of seven participants (70%) identified straight away the veganism aspect of M1 (P1, P5, P9, P13, and P14). A similar group, comprised of P5, P6, P9, P13, and P14, detected users with generic/other food restrictions as being potential users of the application, due to the dietary restrictions feature. P5 and P14 identified sustainability as something important to the user (due to packaging type and sustainable delivery) and P14 mentioned "Health" as something the user is concerned about.

5.3.2 M1 - T2

When explicitly prompted to discuss user values, all participants mentioned veganism and sustainability. All but P6 included users with other generic food restrictions as well. P1, P13, and P14 particularly mentioned packaging type as a concern the user may have. P1, P2, P9, and P14 included "Health" as a user value.

When comparing M1 with the non-value-based MoLIC diagram, many participants decided to apply the answers that were given to the other diagram to M1 as well, claiming that M1 contains most aspects as the other diagram but not the other way around. This was the case with P1, P2, P6, and P14. This group claimed that "Time" is a user value, and, apart from P2, stated that "Ratings" and "Quality" are also important.

Overall, during this task, P1 was the participant who identified the biggest number of values. P1 used the signs that the designer included in the scenes as a way to infer what the user appreciates when ordering food, so items like "Distance" and "Delivery fee" were mentioned at this stage.

5.3.3 M1 - T3

In the third task, participants were asked to grade how appropriate M1 is to the two personas. No values were mentioned at this stage, and participants were mostly silent during the task.

When grading the level of appropriateness of M1 to the personas, P6 claimed that M1 was somewhat appropriate for Sofia (score 4 out of 7), whereas all other participants claimed it to be very appropriate (score 6). P6's reasoning for this was that Sofia feels insecure to order food online because of its quality

and because she cannot be sure whether it is truly vegan or not. Because of this, P6 claimed that no food delivery application would be *very* appropriate for Sofia, who feels insecure.

5.3.4 M1 - T4

For the final task, participants were given a form in which they had to grade how much M1 complies with a list of factors. In this task, P9 explicitly mentioned "Packaging type" and P9 and P14 mentioned "Type of food" as concerns with which M1 does comply.

5.4 M2 - Qualitative results

For M2 (the one with user values present only in the user-designer dialogue), we have participants P3, P4, P7, P8, P10, P11, and P12. The full table of remarks made by participants, broken down by task, can be found in Appendix F.

5.4.1 M2 - T1

In T1, participants were asked to describe the potential users of M2. All participants but one (P4) at first described the user of M2 in a generic manner, such as claiming that users do not have much time to cook or that it is convenient to order food via mobile applications. Some participants also described their own experience with these applications as a way to justify their answer. It was then explained to the participants that they should base their answers on M2 and that they should try to abstract their own bias. This explanation, however, did not result in relevant changes to their answers, apart from P10.

P10, after receiving the explanation above, tried to compare both diagrams that were presented to them (M2 and M4). After looking for differences, they spotted the sustainability concerns related to packaging as well as charity/social concerns. They claimed that the user of M2 really values "Empathy".

P4 managed to identify veganism, generic food restrictions, packaging concerns, charity concerns, and customization options at first glance, with no further explanation from the researcher. These characteristics were not as obvious as the value-based features from M1, as they were present inside the same scenes that are common to all the diagrams. Even after comparing M2 to either M3 and M4, other participants failed to see any difference at this stage.

All participants who received M2 and M3 (P3, P7 and P11) noticed the extra feature in M3 ("Explore discounts"). Because of this comparison, they found the user of M2 to be someone who is not trying to save money.

5.4.2 M2 - T2

After being prompted to spot user values, besides P4 and P10, only P3 was able to identify the values in M2. They mentioned veganism and charity/social concerns. All other participants continued describing the user in a generic manner, with some participants using the same values to describe their other – non-value-based – diagram.

5.4.3 M2 - T3

With the persona in hands, P12 was able to identify user values in M2. These values were: veganism, generic food restrictions, packaging concerns, and charity/social concerns. P10 also noticed that M2 has more a explicit definition of "food category", that includes "gluten free", and "vegan".

5.4.4 M2 - T4

In T4, participants were given a form in which they had to grade how much M1 complies with a list of factors. For the last task, the remaining participants – P7, P8 and P11 – were able to identify user values in M2. All three identified packaging concerns and charity concerns, but only P8 and P11 identified veganism. P11 was more specific and identified other food restrictions and health concerns, sustainability concerns, social concerns, and customization options.

5.5 M3 - Qualitative results

For M3, which was a value-neutral diagram with a slight focus in discounts, we have participants P1, P3, P5, P7, P9, P11 and P13. The full table of remarks made by participants, broken down by task, can be found in Appendix G.

5.5.1 M3 - T1

As expected, when asked to describe the potential users of M3, participants did so in a generic way. All participants but one (M13) described users who

care about prices and discounts. P3, P5, P9, and P13 discussed that the user is interested in the "Convenience" aspect of ordering food online. P3, P5, and P13 mentioned time-saving concerns, and P1, P7, and P9 described the user as having an "Exploratory behavior" due to the features of exploring discounts, restaurants, and dishes.

5.5.2 M3 - T2

After introducing the concept of user values in T2, most participants struggled to understand how this concept would apply to M3. P1 stated that the values mentioned regarding M3 also applied to M1. Two participants – P3 and P13 – did not identify any user values in M3.

5.5.3 M3 - T3

After providing the personas, participants were asked to align each persona with each MoLIC diagram, through a questionnaire. No relevant user values were mentioned at this stage.

P7 and P11 at this stage were still not able to identify user values for M2. When matching the personas to the diagrams, even though no values were mentioned, they decided that Sofia was not an ideal user for M3, for different reasons.

P7 claimed that Sofia was not an ideal user for M3 because she has dietary restrictions and the "Explore discounts" scene may provide options that are not vegan. They explained that she would not explore the discounts, but would rather order from restaurants she is already familiar with. Since Janaína, the more "generic" persona, does not pose restrictions, she would be able to explore the discounts and like them.

In a similar way, P11 found Sofia as someone more concerned with the quality of what she eats, which can be more "gourmet" than most people. Because of that, they decided that M2 was more appropriate for Janaína, who seemed like someone who can appreciate exploring her options and discounts.

5.5.4 M3 - T4

In T4, participants were asked to grade how strongly each MoLIC diagram complies with each factor regarding food delivery applications, through a questionnaire, but no relevant user values were mentioned at this stage.

5.6 M4 - Qualitative results

For M4, the most generic, non-value-based diagram, we have participants P2, P4, P6, P8, P10, P12, and P14. The full table of remarks made by participants broken down by task can be found in Appendix H.

5.6.1 M4 - T1

In T1, participants were asked to describe the potential users of M4. Similarly to M3, participants described M4 in a nonspecific way, using their own experience with food delivery applications as inspiration. All but one participant (P4) described "Convenience" as being important to the user. Most – P2, P8, P10 and P12 – also mentioned "Time".

5.6.2 M4 - T2

With the definition of user values, most participants mentioned "Time" – P2, P6, P8 and P14 – but only two stuck with "Convenience" – P8 and P14.

5.6.3 M4 - T3

No relevant user values were mentioned at this stage.

5.6.4 M4 - T4

P12 mentioned that the user of M4 could be concerned with "Ratings".

5.7 Quantitative results

In this section, we analyze the quantitative results derived from the questionnaires provided in T3 and T4.

In T3, participants were asked to fill out a questionnaire, grading how appropriate each MoLIC diagram is to each persona. The full results can be found in Table 5.2 and are summarized in Figure 5.1. We can see that participants identified that the system described in M1 is more appropriate for Sofia, and Janaína can still use it. M2 is also more appropriate for Sofia, but Janaína may not like it. Both M3 and M4 are more appropriate for Janaína, and Sofia would not like them.

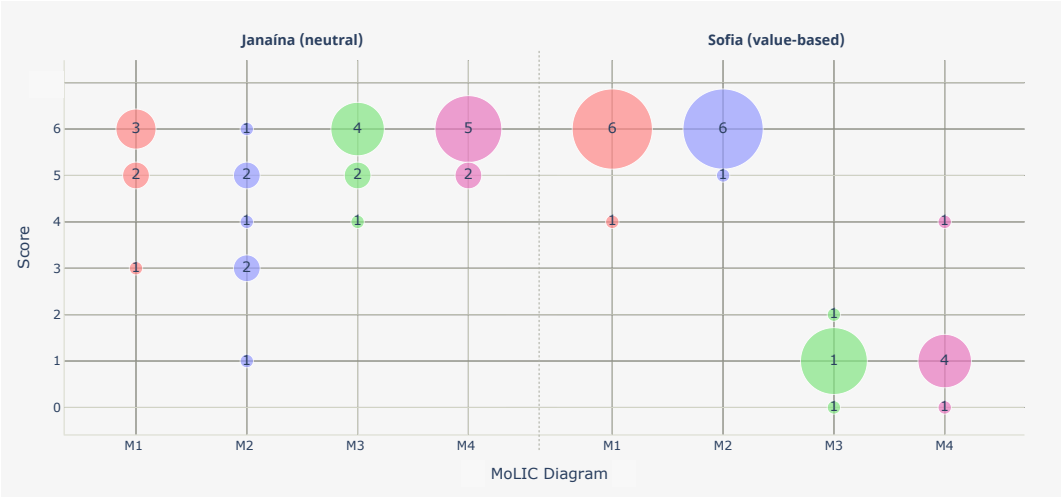

Figure 5.1: Chart displaying how many participants gave each score for each Persona/MoLIC diagram combination

These results are aligned with our expectations. A user with strong values would appreciate more an application designed especially for her, with her values in mind. This does not mean, however, that a user with less defined values cannot use the application: as participants realized, any (generic) user can still perform regular tasks in order to fulfil her goals. A value-based user, though, may struggle with performing simple tasks, as her needs are more specific.

In T4, participants were asked to fill out a similar questionnaire, grading how much each MoLIC diagram complies with each factor/value. The results from this questionnaire can be found in Table 5.3. This list of factors was created based on common values regarding food delivery applications. Some of them, like "Hygiene" and "Data Security", are unrelated to the diagrams.

In the context of Study 2, the most relevant factors are: "Price", which is present in M3, "Familiarity with the restaurant", which is important to the value-based persona, Sofia, "Type of packaging", present in M1 and M2, and "Type of food", somewhat present in M1 and M2 – if we consider Sofia's value of veganism as being related to type of food. Figure 5.2 shows a boxplot chart of these four values.

It is clear that "Type of packaging" received higher scores for M1 and M2. M3 and M4 had no explicit references to what type of packaging was used by the restaurants, hence why most participants graded it 0. Only one participant graded it 1 for M3 (P11) and only one graded it 1 for M4 (P4). When comparing these grades with the participants' own preferences when ordering food online – found in Table 5.4 –, many participants who claimed not to care about type of packaging were still able to see that M1 and M2

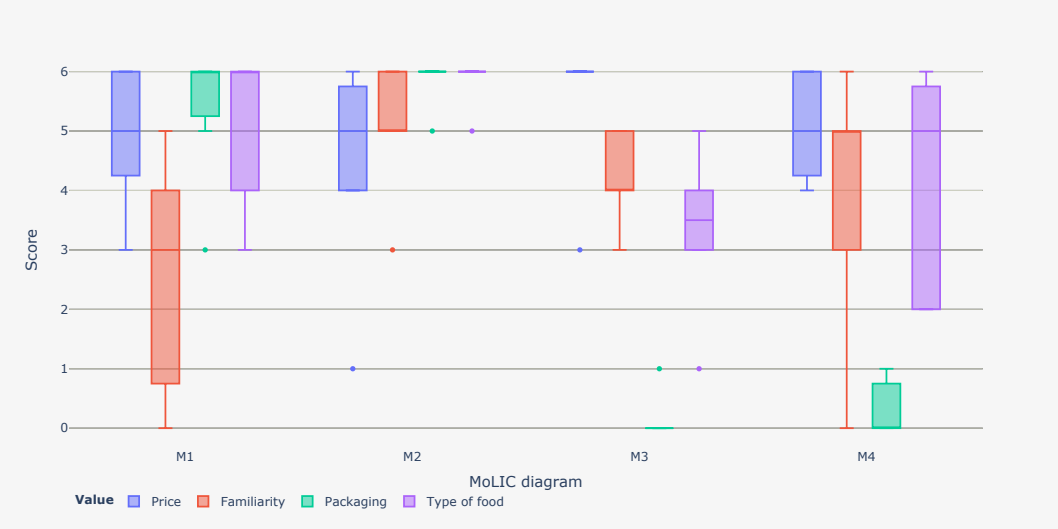

Figure 5.2: Boxplot chart displaying how participants rated Price, Familiarity, Packaging and Type of Food in the MoLIC diagrams

comply with users who do care. For instance, P1, P2, and P11 claimed not to care about packaging type (0 or 1), but still rated M1 or M2 as either 5 or 6 for complying with type of packaging. P9, however, who claimed not to care at all about packaging type (0), rated M1 as 3 (middle ground) for type of packaging. This was the lowest score for either M1 or M2 regarding packaging type and was not verbally justified.

Table 5.2: Scores given by participants regarding how well each diagram is aligned with each persona

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
M1 - Janaína	6	6			5	6			3				5	3
M1 - Sofia	6	6			6	4			6				6	6
M2 - Janaína			5	6			3	5		3	1	4		
M2 - Sofia			6	6			5	6		6	6	6		
M3 - Janaína	6		5		6		4		6		6		5	
M3 - Sofia	1		1		1		1		2		1		0	
M4 - Janaína		6		6		6		6		5		5		6
M4 - Sofia		1		1		0		4		1		3		1

Table 5.3: Scores given by participants regarding how well each diagram satisfies each factor/value

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
M1 - Price	5	6			4	6			5				3	6
M1 - Familiarity with the restaurant	5	4			3	3			0				4	0
M1 - Hygiene	5	2			3	2			0				5	3
M1 - Type of packaging	5	6			6	6			3				6	6
M1 - Delivery time	6	6			6	5			6				5	6
M1 - Type of food	6	4			-	6			3				6	6
M1 - Data security	5	1			2	0			0				3	0
M2 - Price			4	5			5	6		6	1	4		
M2 - Familiarity with the restaurant			6	6			5	6		5	5	3		
M2 - Hygiene			5	3			4	4		6	1	2		
M2 - Type of packaging			6	6			5	6		6	6	6		
M2 - Delivery time			6	5			5	6		4	5	6		
M2 - Type of food			6	6			6	6		6	6	5		
M2 - Data security			5	3			0	6		3	1	3		
M3 - Price	6		6		6		6		6		6		3	
M3 - Familiarity with the restaurant	5		4		3		5		5		5		4	
M3 - Hygiene	5		5		3		4		1		1		3	
M3 - Type of packaging	0		0		0		0		0		1		0	
M3 - Delivery time	6		6		6		4		6		5		5	
M3 - Type of food	6		4		-		1		3		3		5	
M3 - Data security	5		5		3		0		0		1		3	
M4 - Price		6		5		4		4		6		6		6
M4 - Familiarity with the restaurant		4		5		3		6		5		3		0
M4 - Hygiene		0		3		0		4		4		2		2
M4 - Type of packaging		0		1		0		0		0		0		0
M4 - Delivery time		6		5		5		6		4		6		6
M4 - Type of food		0		2		6		6		5		2		5
M4 - Data security		1		3		0		6		3		3		0

Table 5.4: Grades given by participants regarding much they care about each factor/value when ordering food in mobile applications

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
Price	6	5	4	5	6	6	4	6	3	6	5	5	5	4
Familiarity with the restaurant	5	4	6	6	4	4	6	5	5	6	2	5	1	6
Hygiene	5	0	5	6	6	6	6	6	6	6	1	6	4	6
Type of packaging	1	0	6	4	5	4	4	2	0	3	1	4	4	3
Delivery time	4	6	5	3	5	6	4	6	5	5	4	5	1	5
Type of food	4	0	6	5	5	5	2	5	6	5	6	6	6	6
Data security	1	0	6	4	5	6	5	6	6	6	3	3	5	6

5.8 Discussion - Study 2

Both M1 and M2 had most of their values detected by the participants. The biggest difference between the two is that M1 has new scenes, *i.e.*, the interaction flow is different. The addition of scenes was more noticeable to participants than the addition of new signs inside preexisting scenes. Because of that, participants who analyzed M1 were quicker to identify user values – which happened mainly during T1 and also T2 – and needed fewer prompts than those who analyzed M2.

This raises new questions: is it an issue that people sometimes need to be prompted to recognize user values in design artifacts? Do user values need to be highlighted more explicitly in design artifacts so as to be recognized and communicated more easily? In order to answer these new questions, further research is necessary. Many participants in the study first analyzed the diagrams in a superficial way, whereas in a real scenario of software development, it is expected that designers or software engineers do a more thorough reading.

A possible next step in this research would be to actually implement the system described by the diagrams, *i.e.*, make the scenario more concrete. This would imply that participants should read the diagrams more carefully, focusing more on the details. Ideally, the resulting software, which was to be constructed based on the evaluation of the interaction diagrams, would contain features rooted in user values. If this were the case, then it would be fair to say that the value information communicated through the MoLIC diagrams is enough to meet the users' values.

Another question that we need to address is the fact that many participants compared both diagrams they received when answering the questions. We did change the order in which we provided the diagrams (either value-based then non-value-based, or non-value-based then value-based) during the study, but ultimately many participants ended up comparing and looking for differences between the two. This happened mostly with participants who got M2, because of its subtle changes. Future research can involve only one diagram per participant, as a way to avoid comparisons. We did not try this approach as we had a limited number of participants, and one diagram per person would weaken our results.

Regardless, as seen in Figures 5.1 and 5.2, M1 and M2 embed values, such as "Type of packaging", which are important to Sofia. Keep in mind that M1 and M2 were created specifically for her based on the results from Study 1. Participants from Study 2 rated M1 and M2 as being highly appropriate for

her, whereas M3 and M4 as being inappropriate because of her specific needs regarding food delivery applications. Janaína, with no explicit user values, seemed like a better fit for M3 and M4, which were also made for her needs. She can still, however, use M1 and M2, with most participants claiming it to be either appropriate or neutral for her.

5.8.1 Answering RQ2b

The Research Question 2b is: "If a design artifact embeds values, will designers recognize those values?" This question is related mainly to M1 and M2, which were the value-based MoLIC diagrams representing the design artifacts.

We believe that yes, designers do recognize these values. Sections 5.3 and 5.4 go over how participants imagined M1 and M2's potential users and what their values might be, also discussing the moment of their remarks. Some participants needed prompts to discuss user values specifically, while others needed to focus and read the diagram more carefully.

As discussed previously, participants often performed comparisons between the non-value-based and the value-based diagram. Despite this comparison, participants did manage to identify that the difference could represent something important to the user whose needs and preferences led to the diagram they were analyzing. The fact that the participants were sensitive enough to acknowledge that M1 and M2's user was vegan, cared about the environment, and supported social causes is a positive outcome. It leads us to believe that user values can be respected and satisfied in a design and software engineering context.

Another interesting aspect of Study 2 was that some participants appeared to feel empathetic towards people with dietary restrictions. A few participants claimed not to have any restrictions themselves, but that it was a good thing that M1 included a feature for those who do have. In a field such as Computer Science, where professionals may struggle with understanding the needs and values of the users of the systems they are creating, it is refreshing to see young professionals empathize with other people's values.

6
Conclusion

In this work, we explored the integration of user values with design rationale using QOC and how it impacts a design artifact, namely MoLIC diagrams. We conducted two separate studies to answer the following research questions:

RQ1: How can value sensitive design be integrated into design rationale?

RQ2: How can design artifacts reflect the design rationale taking into account value sensitive design?

RQ2a: Given a design rationale which explicitly references values, do designers embed those values in the design artifacts, explicitly or implicitly?

RQ2b: If a design artifact embeds values, will designers recognize those values?

In the first study, participants were asked to record their design rationale regarding new features for a food delivery application based on a persona, which was either value-based or non-value-based. On one hand, we found that those with the value-based persona made decisions based mainly on the persona's values and preferences. The participants noted on the design rationale document when user values impacted their decisions, which was frequent. On the other hand, the participants who received the more "generic" persona made decisions mainly based on their own biases and experience with food delivery applications.

For the first study, participants were asked to edit a simple MoLIC diagram of a food delivery application taking into account the decisions they made during their design rationale process. Participants did not explicitly highlight user values into the MoLIC diagrams, but we believe certain changes were clearly inspired by the value-based persona (*e.g.*, the addition of veganism-related features). We found participants who received the value-based persona did come up with features more targeted to their persona, whereas the other participants included basic, generic features.

Although we were able to see these differences, we conducted the second study to assess whether other designers, oblivious to the first study, had the

same perception. We found that, indeed, the participants could perceive user values in the MoLIC diagrams, *i.e.*, even though the values were not explicitly embedded, designers could apprehend the user's values.

Although both groups used personas as a method for depicting users, it is clear that a well thought-out persona, such as Group V's, results in a richer design artifact. Group N's persona was too bland and did not carry strong values like Group V's, so designers tended to use their own experience as a guide. Therefore, it is not enough to use personas to guide the design process – they should be carefully crafted and should truly reflect the users' values, or else the resulting software may not be appropriate for them and may be created based on the designers' own values.

A possibility of future research can be an adaptation of the Questions, Options, and Criteria method to incorporate user values in the decision-making process. The participants of the first study did so by circling or underlining options motivated by the persona, but there was no formal way to do it. If the QOC formally included user values, perhaps this could have been done in a more systematic way.

There are many other design rationale notations other than the QOC. New studies can be conducted with other methods, such as IBIS (Kunz and Rittel, 1970), with the aim of finding whether the integration of user values with DR can happen in a more efficient way. Another possibility is the creation of a new DR method that explicitly includes the evaluation of user values. DR methods are used as a decision-making tool, and software design deals with countless instances of decision-making regarding users and other factors. Therefore, a new method for recording the rationale behind these decisions with the final user in mind can be of great benefit to the usability and user experience of a software artifact.

Similarly, the MoLIC notation could be updated to explicitly support user values. Perhaps the reason why participants of the first study failed to highlight the influence of user values in the MoLIC diagram is that the notation itself is not prepared to handle these sort of accounts. Because SemEng, the theory underlying MoLIC, does not *explicitly* tackle user values, the resulting language does not emphasize the importance of user values either. If MoLIC is updated to reflect user values, it is possible that the final software product derived from the modelling process will become more appropriate for the user and their needs, preferences and values.

A different approach can be the creation of a new modelling language designed specifically for user values. It can include how values are connected to each other in a specific context and how different personas feel about each

value. That way, designers can have a clear vision of the contextual impact that user values can have in the software in question, and understand each persona's needs. This clear view may also facilitate the DR process and help with understanding the importance of each value, which can be used as criteria.

In our studies, we did not use the SemEng metacommunication template as a tool, but it is an intriguing idea to embed user values into the template and then use it as a support tool for building the MoLIC diagram. By relying on a clear understating of how the user will use the system, and his/her values, we could generate richer diagrams, with more mentions to user values. Hence, future work can involve looking into the metacommunication template and incorporating user values into it.

With this work, we try to showcase the importance of taking into account user values in the design process. The results of participants who dealt with and discussed user values were much richer than those who received a non-value-based persona. Therefore, new ways or methods of including user values in the design process can be developed. Value Sensitive Design provides a good array of methods, but can be overwhelming to software engineers. A middle ground between Value Sensitive Design and Values-First Software Engineering could be the way to go. New methods should be simple and easily implemented by software engineers and designers, in order to ensure the resulting software to be as appropriate as possible for the user and her own values. As put by Davis and Nathan (2015), we too believe that, in the future, user values will be highly appreciated in technology design.

Bibliography

Barbosa, S. D. J. and de Paula, M. G. (2003). Designing and Evaluating Interaction as Conversation: A Modeling Language Based on Semiotic Engineering. In Jorge, J. A., Jardim Nunes, N., and Falcão e Cunha, J., editors, *Interactive Systems. Design, Specification, and Verification*, Lecture Notes in Computer Science, pages 16–33. Springer Berlin Heidelberg.

Borning, A., Friedman, B., Davis, J., and Lin, P. (2005). Informing Public Deliberation: Value Sensitive Design of Indicators for a Large-Scale Urban Simulation. In Gellersen, H., Schmidt, K., Beaudouin-Lafon, M., and Mackay, W., editors, *ECSCW 2005*, pages 449–468. Springer Netherlands.

Carroll, J. M. (1997). Human-computer interaction. *International Journal of Human-Computer Studies*, 46(4):501–522.

Carroll, J. M. and Rosson, M. B. (2003). Chapter 15 - design rationale as theory. In Carroll, J. M., editor, *HCI Models, Theories, and Frameworks*, Interactive Technologies, pages 431 – 461. Morgan Kaufmann, San Francisco.

Conklin, J. and Begeman, M. L. (1987). gIBIS: A Hypertext Tool for Team Design Deliberation. In *Proceedings of the ACM Conference on Hypertext*, HYPERTEXT '87, pages 247–251, New York, NY, USA. ACM. event-place: Chapel Hill, North Carolina, USA.

Conklin, J. and Begeman, M. L. (1988). gIBIS: A Hypertext Tool for Exploratory Policy Discussion. *ACM Trans. Inf. Syst.*, 6(4):303–331.

Davis, J. and Nathan, L. P. (2015). Value Sensitive Design: Applications, Adaptations, and Critiques. In van den Hoven, J., Vermaas, P. E., and van de Poel, I., editors, *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, pages 11–40. Springer Netherlands, Dordrecht.

de Souza, C. S. (2005a). Semiotic engineering: bringing designers and users together at interaction time. *Interacting with Computers*, 17(3):317–341.

de Souza, C. S. (2005b). *The Semiotic Engineering of Human-Computer Interaction (Acting with Technology)*. The MIT Press.

de Souza, C. S. and Leitão, C. F. (2009). Semiotic Engineering Methods for Scientific Research in HCI. *Synthesis Lectures on Human-Centered Informatics*, 2(1):1–122.

de Souza, C. S., Leitão, C. F., Prates, R. O., and da Silva, E. J. (2006). The semiotic inspection method. In *Proceedings of VII Brazilian symposium on Human factors in computing systems*, IHC '06, pages 148–157, Natal, RN, Brazil. Association for Computing Machinery.

Denning, T., Kramer, D. B., Friedman, B., Reynolds, M. R., Gill, B., and Kohno, T. (2014). CPS: beyond usability: applying value sensitive design based methods to investigate domain characteristics for security for implantable cardiac devices. In *Proceedings of the 30th Annual Computer Security Applications Conference*, ACSAC '14, pages 426–435, New Orleans, Louisiana, USA. Association for Computing Machinery.

Detweiler, C. and Harbers, M. (2014). Value Stories: Putting Human Values into Requirements Engineering. In *Proceedings of Workshop on Creativity in Requirements Engineering (CreaRE)*, Essen, Germany.

Ferrario, M. A., Simm, W., Forshaw, S., Gradinar, A., Smith, M. T., and Smith, I. (2016). Values-first SE: Research Principles in Practice. In *Proceedings of the 38th International Conference on Software Engineering Companion*, ICSE '16, pages 553–562, New York, NY, USA. ACM. event-place: Austin, Texas.

Fischer, G. (2001). User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, 11(1):65–86.

Friedman, B. (1997). Social Judgments and technological innovation: Adolescents' understanding of property, privacy, and electronic information. *Computers in Human Behavior*, 13(3):327–351.

Friedman, B., Hendry, D. G., and Borning, A. (2017). A Survey of Value Sensitive Design Methods. *Found. Trends Hum.-Comput. Interact.*, 11(2):63–125.

Friedman, B. and Kahn, P. H. (2002). Human values, ethics, and design. In *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, pages 1177–1201. L. Erlbaum Associates Inc., USA.

Friedman, B., Kahn, P. H., and Borning, A. (2008). Value Sensitive Design and Information Systems. In *The Handbook of Information and Computer Ethics*, pages 69–101. John Wiley & Sons, Ltd.

Friedman, B., Kahn, P. H., Hagman, J., Severson, R. L., and Gill, B. (2009). The Watcher and the Watched: Social Judgments about Privacy in a Public Place. In Harrison, S., editor, *Media Space 20 + Years of Mediated Life*, Computer Supported Cooperative Work, pages 145–176. Springer, London.

Harbers, M., Detweiler, C., and Neerincx, M. A. (2015). Embedding Stakeholder Values in the Requirements Engineering Process. In Fricker, S. A. and Schneider, K., editors, *Requirements Engineering: Foundation for Software Quality*, Lecture Notes in Computer Science, pages 318–332, Cham. Springer International Publishing.

Jarczyk, A. P. J., Loffler, P., and Shipmann, F. M. (1992). Design rationale for software engineering: a survey. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, volume ii, pages 577–586 vol.2.

Kahn Jr., P. H., Friedman, B., Pérez-Granados, D. R., and Freier, N. G. (2006). Robotic pets in the lives of preschool children. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 7(3):405–436.

Kujala, S. and Väänänen-Vainio-Mattila, K. (2009). Value of Information Systems and Products: Understanding the Users' Perspective and Values. *Journal of Information Technology Theory and Application (JITTA)*, 9(4).

Kunz, W. and Rittel, H. W. J. (1970). *Issues as Elements of Information Systems*. Institute of Urban and Regional Development, University of California.

Lee, J. and Lai, K.-Y. (1991). What's in Design Rationale? *Hum.-Comput. Interact.*, 6(3):251–280.

López-Jaquero, V. and Montero, F. (2007). Comprehensive Task and Dialog Modelling. In Jacko, J. A., editor, *Human-Computer Interaction. Interaction Design and Usability*, Lecture Notes in Computer Science, pages 1149–1158. Springer Berlin Heidelberg.

MacLean, A., Young, R. M., Bellotti, V. M. E., and Moran, T. P. (1991). Questions, Options, and Criteria: Elements of Design Space Analysis. *Human–Computer Interaction*, 6(3-4):201–250.

Marques, A. B., Barbosa, S. D. J., and Conte, T. (2016). A Comparative Evaluation of Interaction Models for the Design of Interactive Systems. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*,

SAC '16, pages 173–180, New York, NY, USA. ACM. event-place: Pisa, Italy.

Miller, J. K., Friedman, B., Jancke, G., and Gill, B. (2007). Value Tensions in Design: The Value Sensitive Design, Development, and Appropriation of a Corporation's Groupware System. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work*, GROUP '07, pages 281–290, New York, NY, USA. ACM. event-place: Sanibel Island, Florida, USA.

Mougouei, D. (2020). Engineering Human Values in Software through Value Programming. In *arXiv:2003.04477 [cs]*.

Mougouei, D., Perera, H., Hussain, W., Shams, R., and Whittle, J. (2018). Operationalizing human values in software: a research roadmap. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2018, pages 780–784, New York, NY, USA. Association for Computing Machinery.

Noble, D. and Rittel, H. W. J. (1988). Issue-Based Information Systems for Design. In *Computing in Design Education [ACADIA Conference Proceedings] Ann Arbor (Michigan / USA) 28-30 October 1988, pp. 275-286*. CUMINCAD.

Prates, R. O., de Souza, C. S., and Barbosa, S. D. J. (2000). Methods and tools: a method for evaluating the communicability of user interfaces. *Interactions*, 7(1):31–38.

Regli, W. C., Hu, X., Atwood, M., and Sun, W. (2000). A Survey of Design Rationale Systems: Approaches, Representation, Capture and Retrieval. *Engineering with Computers*, 16(3):209–235.

Salgado, L. C. d. C., de Souza, C. S., and Leitão, C. F. (2011). Using Metaphors to Explore Cultural Perspectives in Cross-Cultural Design. In Rau, P. L. P., editor, *Internationalization, Design and Global Development*, Lecture Notes in Computer Science, pages 94–103, Berlin, Heidelberg. Springer.

Salgado, L. C. d. C., Souza, C. S. d., and Leitao, C. F. (2009). Conceptual Metaphors for Designing Multi-cultural Applications. In *2009 Latin American Web Congress*, pages 105–111. ISSN: null.

Schwartz, S. H. and Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, 53(3):550–562.

Sutcliffe, A. (2013). Bridging users' values and requirements to architecture. In *2013 3rd International Workshop on the Twin Peaks of Requirements and Architecture (TwinPeaks)*, pages 15–21.

Tang, A., Babar, M. A., Gorton, I., and Han, J. (2006). A survey of architecture design rationale. *Journal of Systems and Software*, 79(12):1792–1804.

Thew, S. and Sutcliffe, A. (2008). Investigating the Role of 'Soft Issues' in the RE Process. In *2008 16th IEEE International Requirements Engineering Conference*, pages 63–66. ISSN: 2332-6441.

Tisdale, J. R. (1961). *Psychological value theory and research: 1930-1960*. Thesis (Ph.D.), Boston University.

Winter, E., Forshaw, S., and Ferrario, M. A. (2018). Measuring Human Values in Software Engineering. In *Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, ESEM '18, pages 48:1–48:4, New York, NY, USA. ACM. event-place: Oulu, Finland.

Winter, E., Forshaw, S., Hunt, L., and Ferrario, M. A. (2019a). Advancing the Study of Human Values in Software Engineering. In *Proceedings of the 12th International Workshop on Cooperative and Human Aspects of Software Engineering*, CHASE '19, pages 19–26, Piscataway, NJ, USA. IEEE Press. event-place: Montreal, Quebec, Canada.

Winter, E., Forshaw, S., Hunt, L., and Ferrario, M. A. (2019b). Towards a Systematic Study of Values in SE: Tools for Industry and Education. In *Proceedings of the 41st International Conference on Software Engineering: New Ideas and Emerging Results*, ICSE-NIER '19, pages 61–64, Piscataway, NJ, USA. IEEE Press. event-place: Montreal, Quebec, Canada.

Yue, G., Liu, J., and Hou, Y. (2018). Design Rationale Knowledge Management: A Survey. In Luo, Y., editor, *Cooperative Design, Visualization, and Engineering*, Lecture Notes in Computer Science, pages 245–253. Springer International Publishing.

A
Participant Profile Questionnaire

Seu perfil

Qual é a sua faixa etária? *

◯ até 20
◯ 21-30
◯ 31-40
◯ 41-50
◯ 51-60
◯ acima de 60

Conhecimento sobre aplicativos de *delivery* de comida

Com que frequência você... *

...pede comida em aplicativos:

◯ 0 - Nunca
◯ 1 - Todo ano
◯ 2 - Todo semestre
◯ 3 - Todo trimestre
◯ 4 - Todo mês
◯ 5 - Toda semana
◯ 6 - Todo dia

Marque fatores que influenciam o processo de pedir comida em aplicativos: *

Preço

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Familiaridade com o restaurante

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Higiene

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Tipo de embalagem usada

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Tempo de entrega

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Tipo de comida

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

Segurança dos dados

Não me importo - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Me importo muito

B
Study 1 - MoLIC diagrams edited by Group N and Group V

B.1
Group N

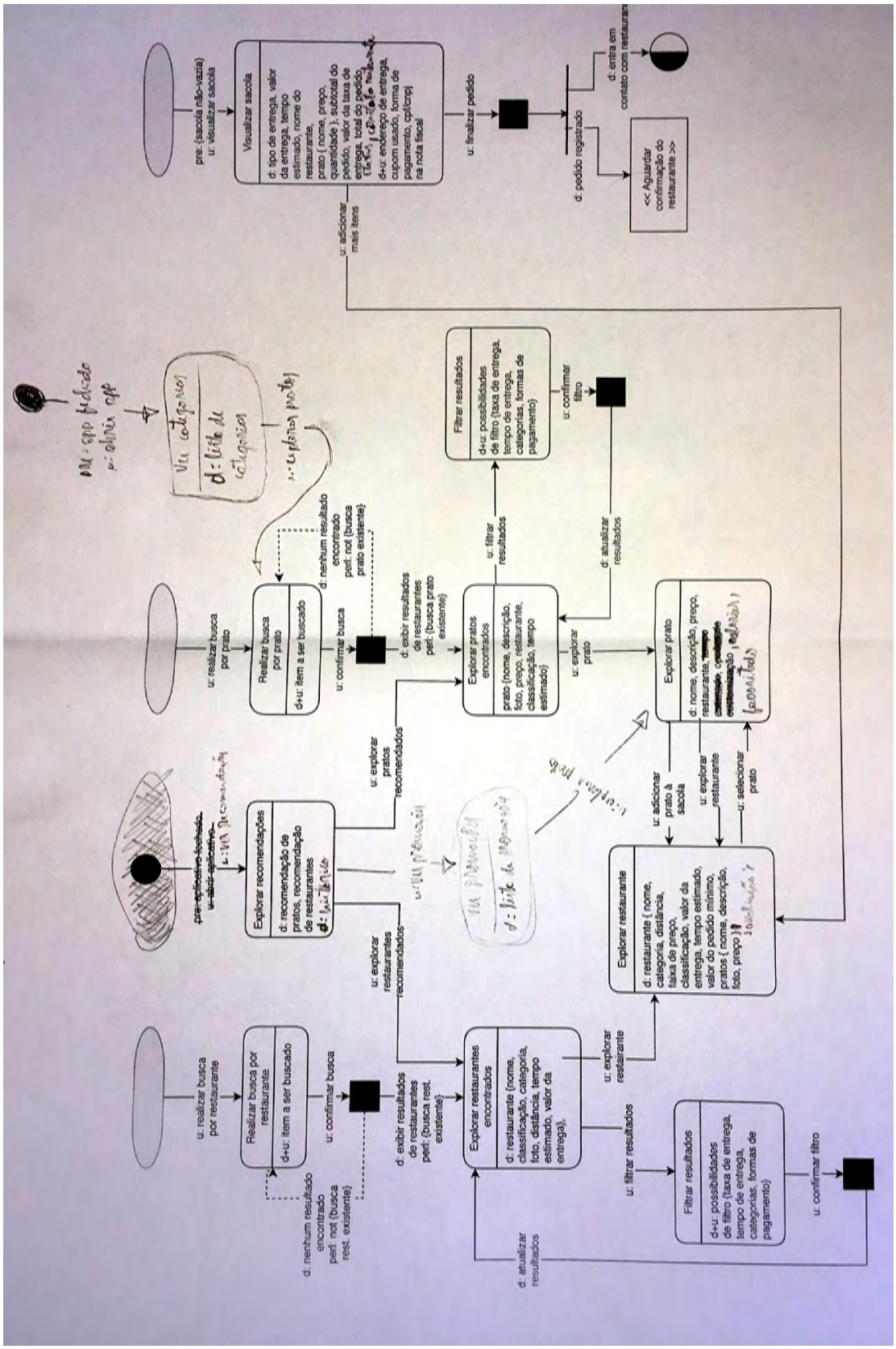

Figure B.1: N1's MoLIC Diagram

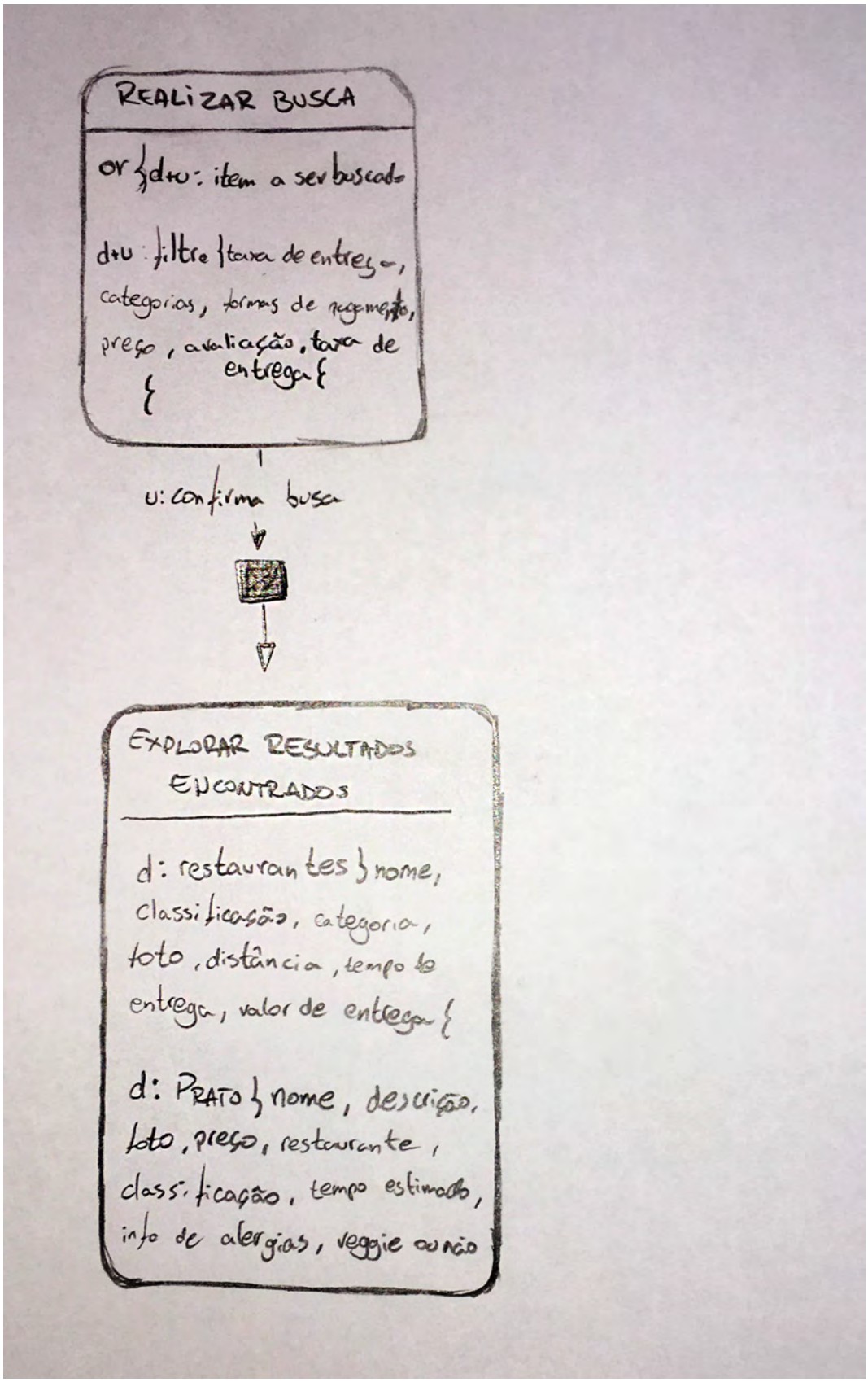

Figure B.2: N2's MoLIC Diagram

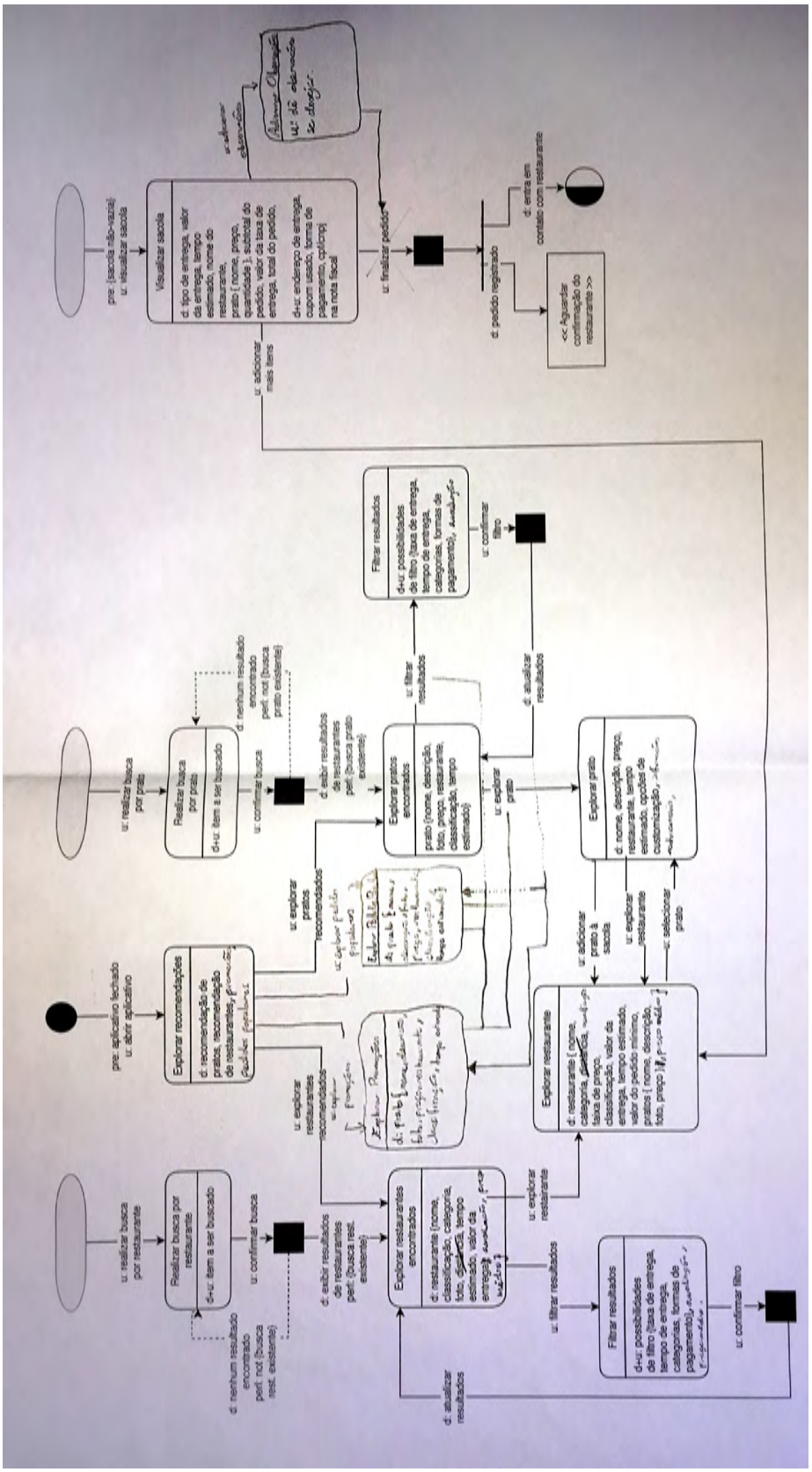

Figure B.3: N3's MoLIC Diagram

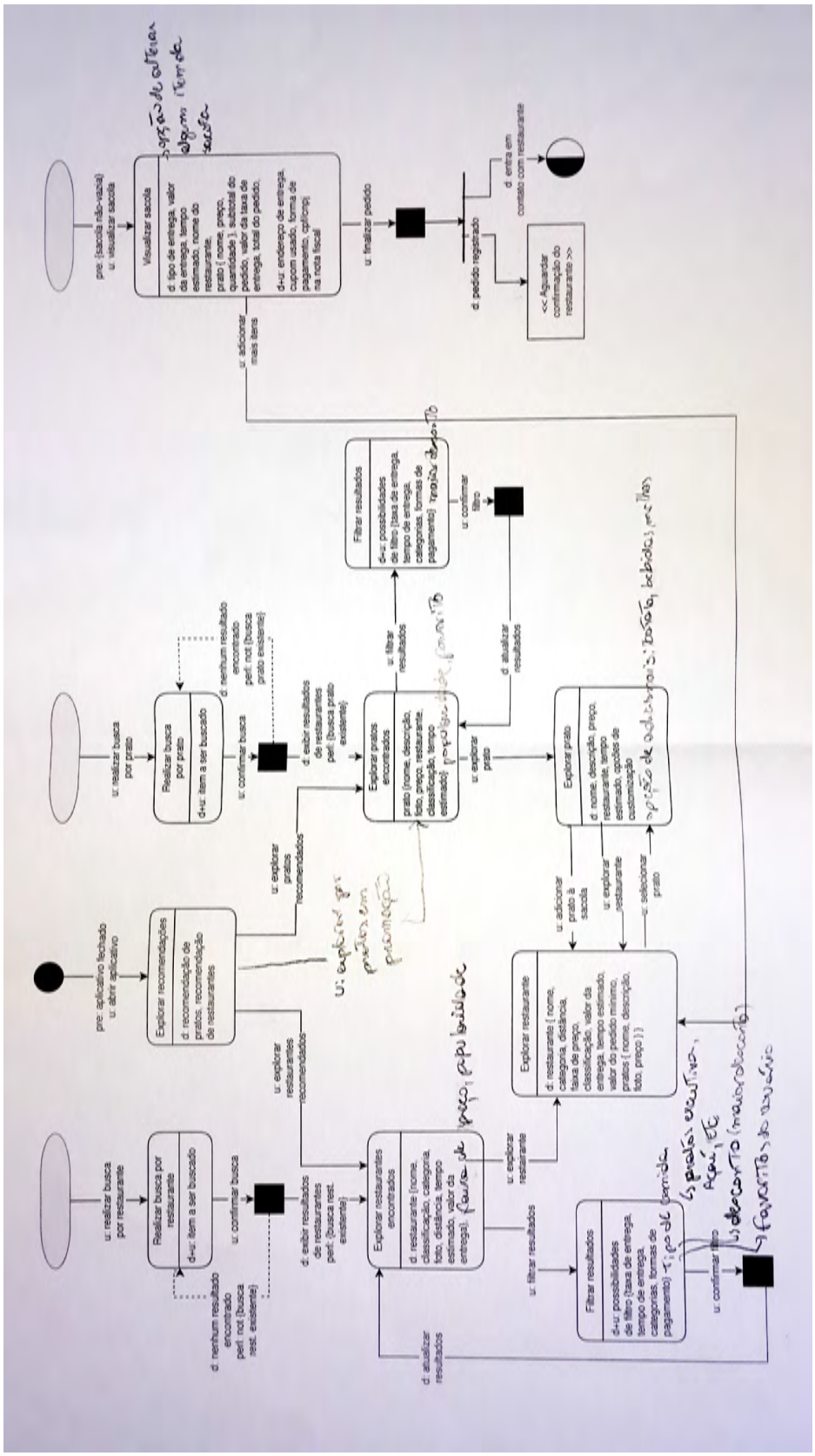

Figure B.4: N4's MoLIC Diagram

B.2
Group V

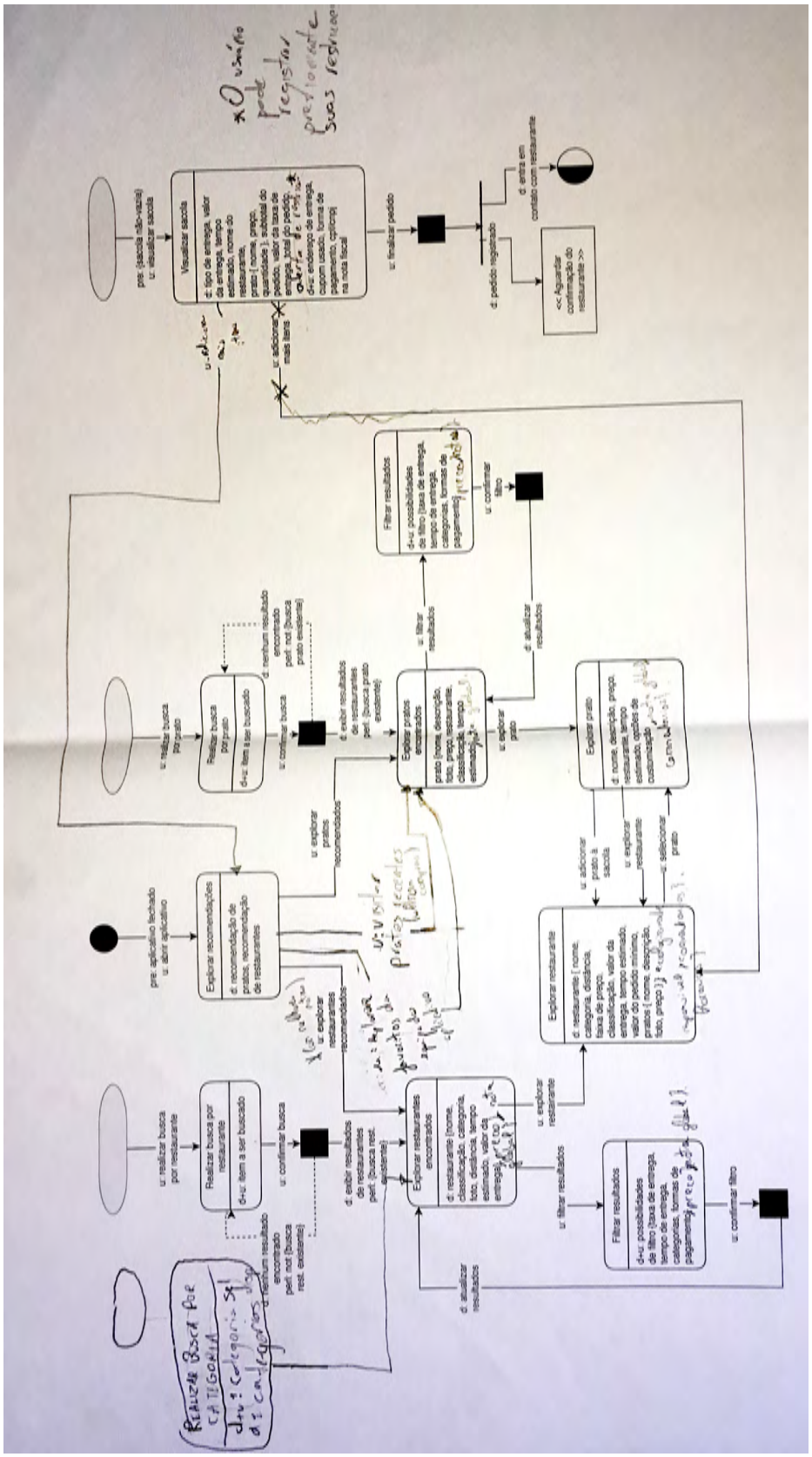

Figure B.5: V1's MoLIC Diagram

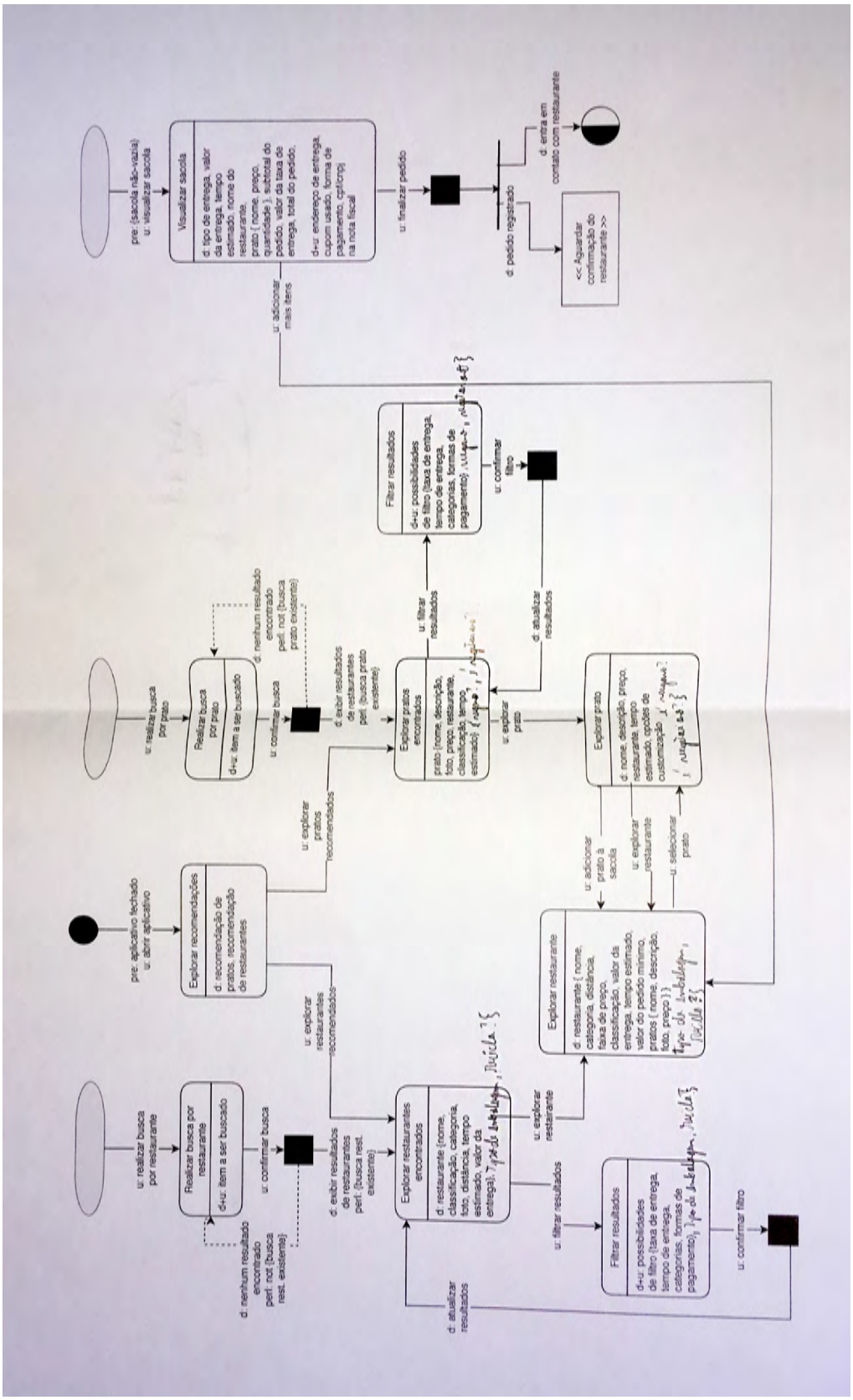

Figure B.6: V2's MoLIC Diagram

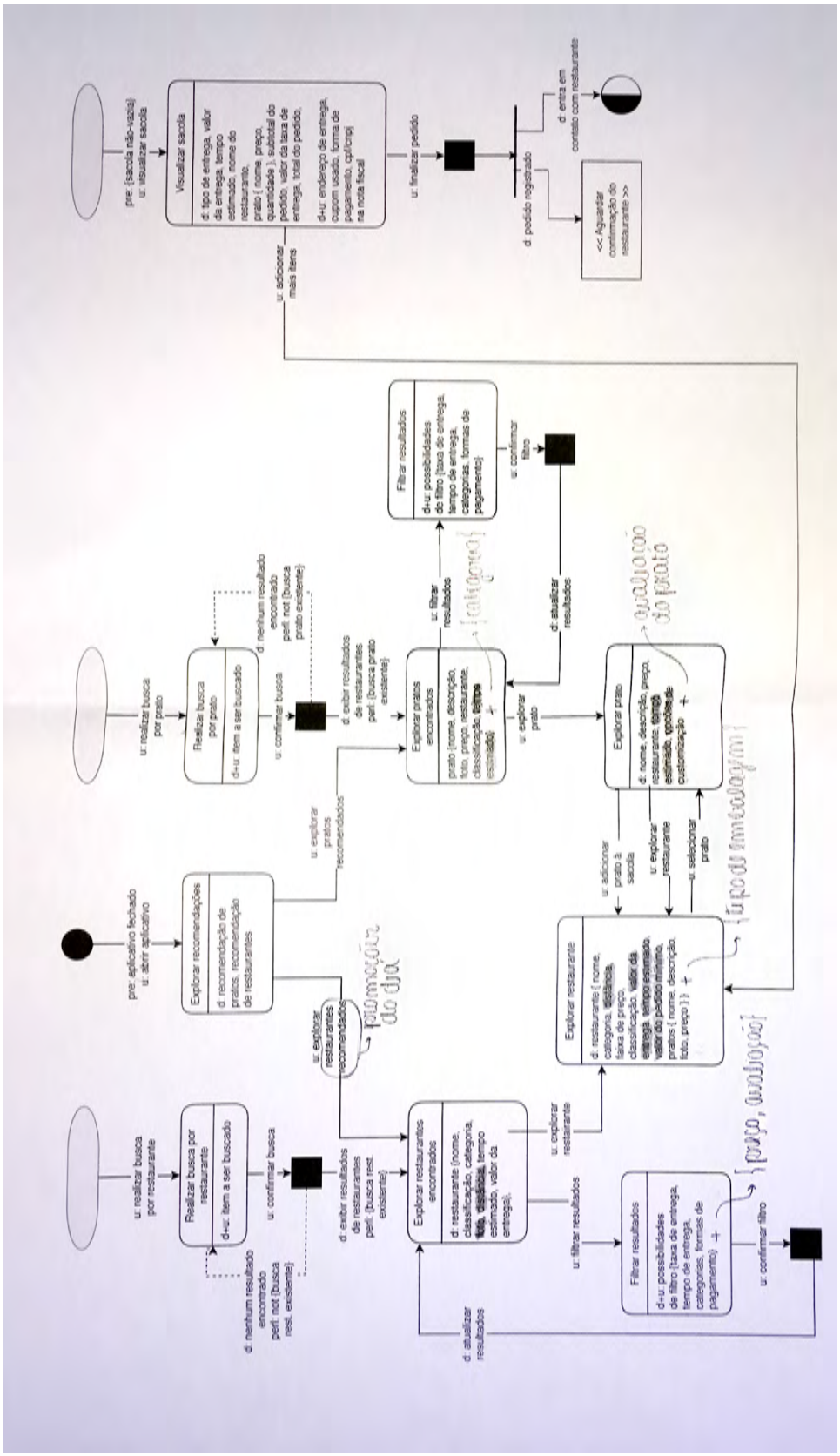

Figure B.7: V3's MoLIC Diagram

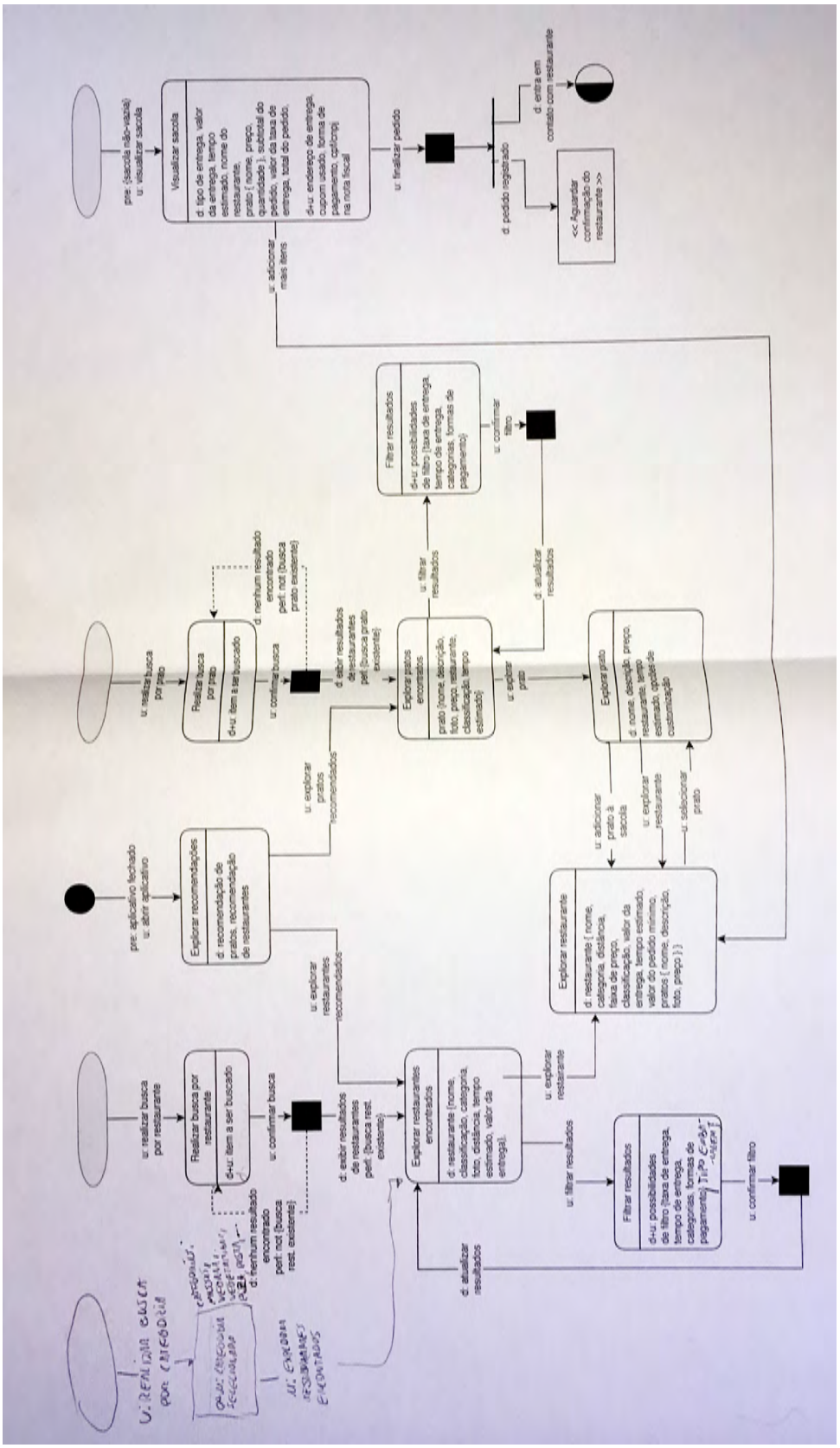

Figure B.8: V4's MoLIC Diagram

C
Study 2 - MoLIC diagrams

C.1
Value-based MoLIC diagrams

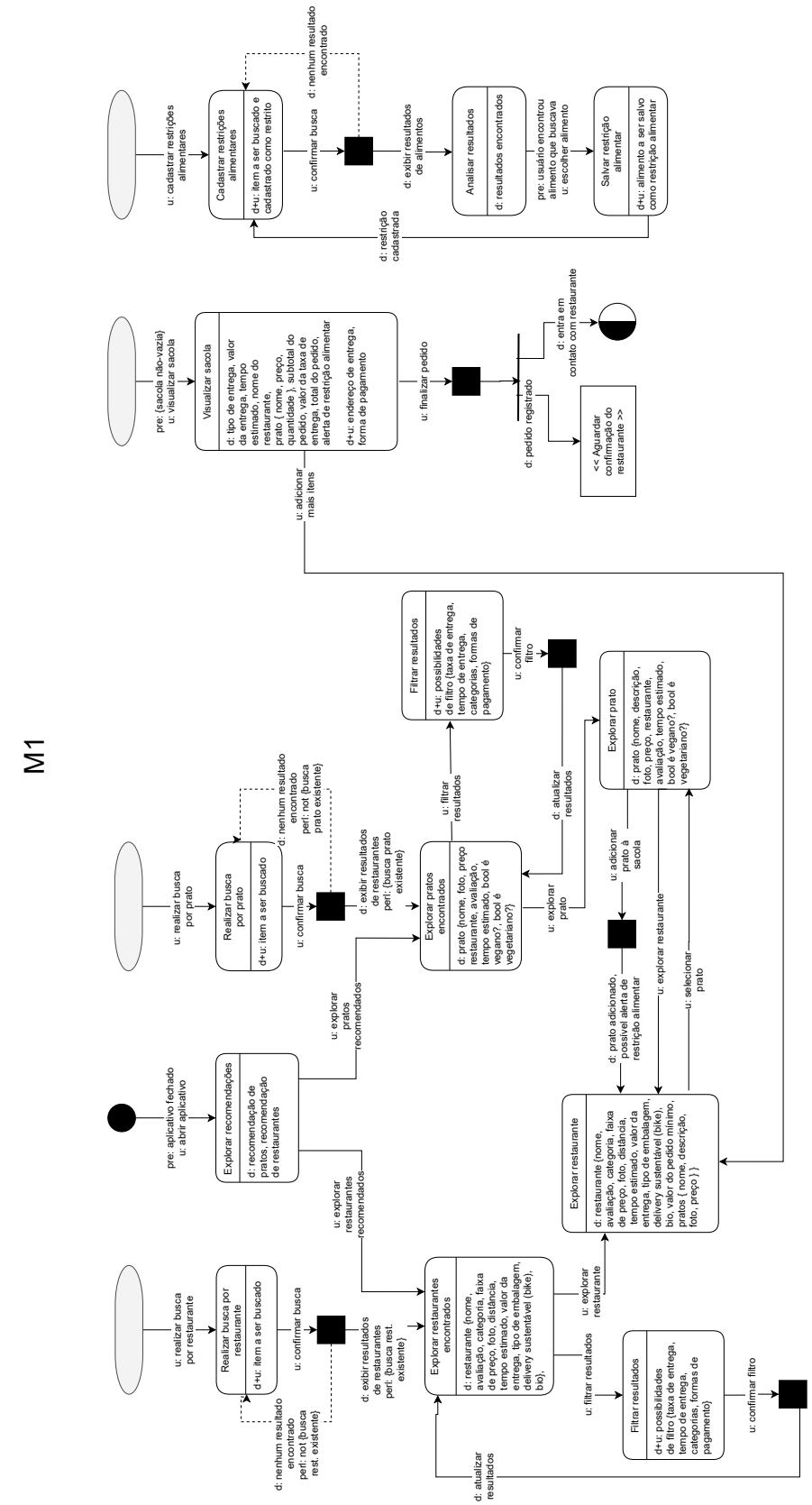

Figure C.1: M1

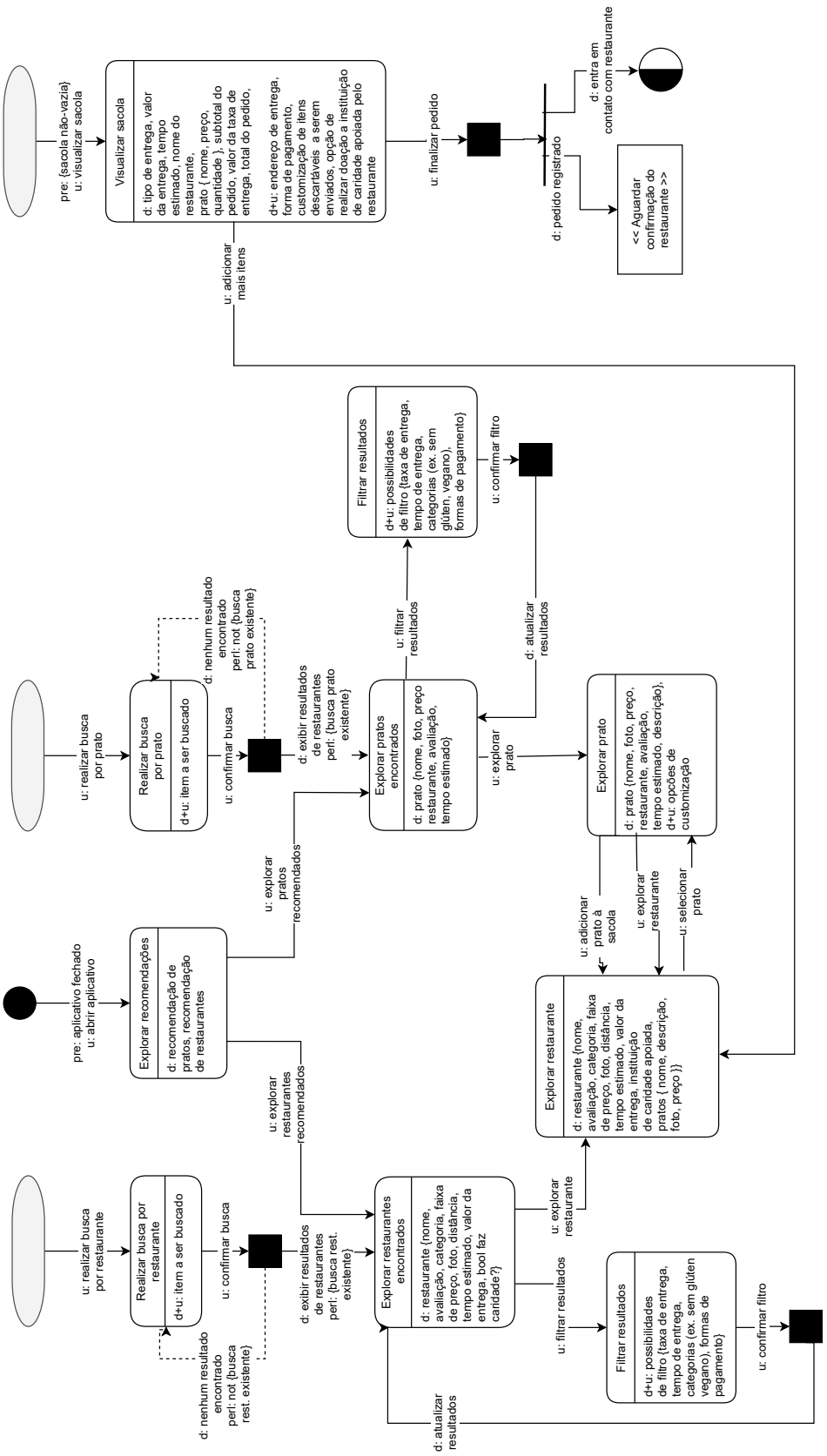

Figure C.2: M2

C.2
Non-value-based MoLIC diagrams

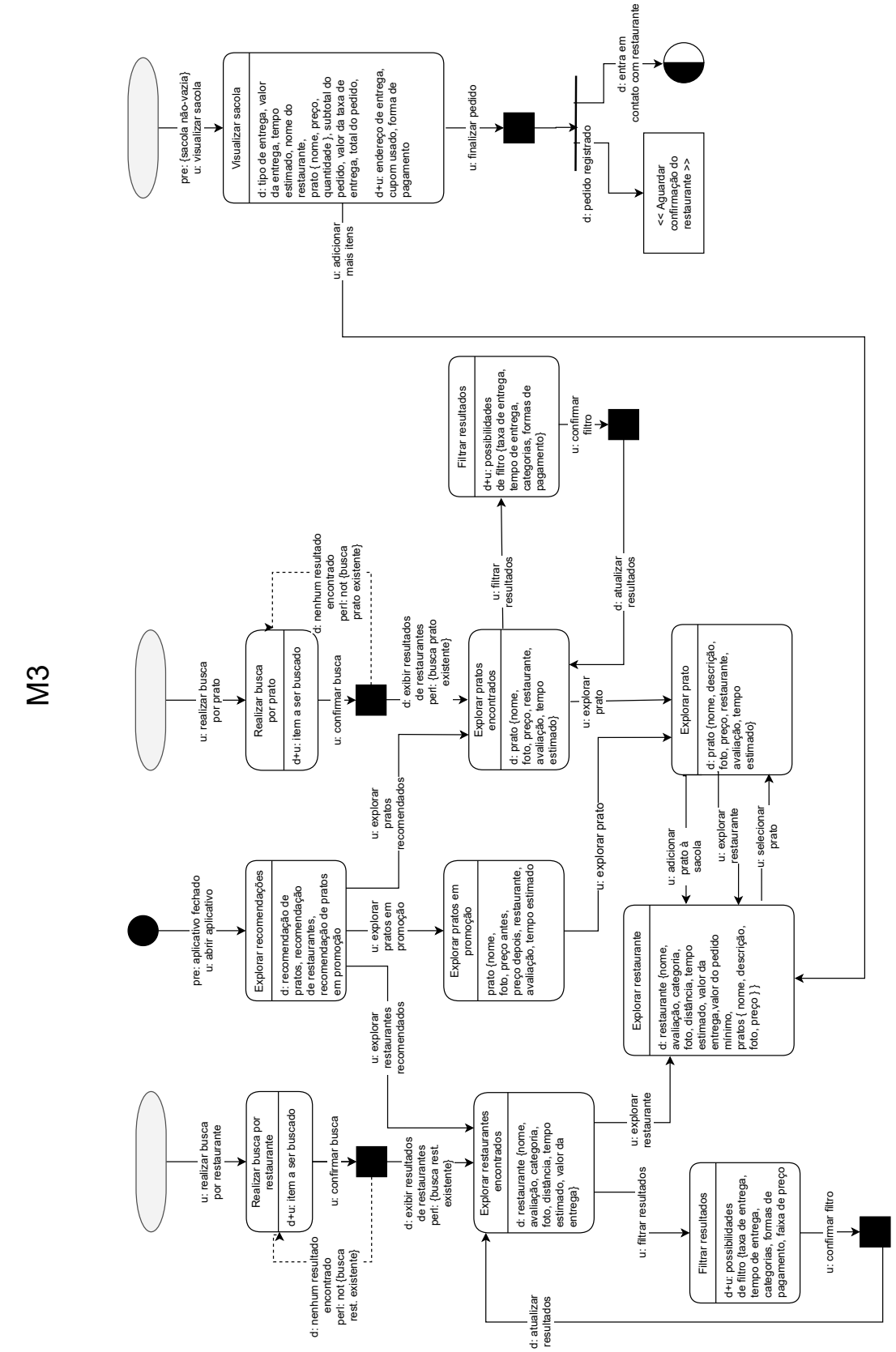

Figure C.3: M3

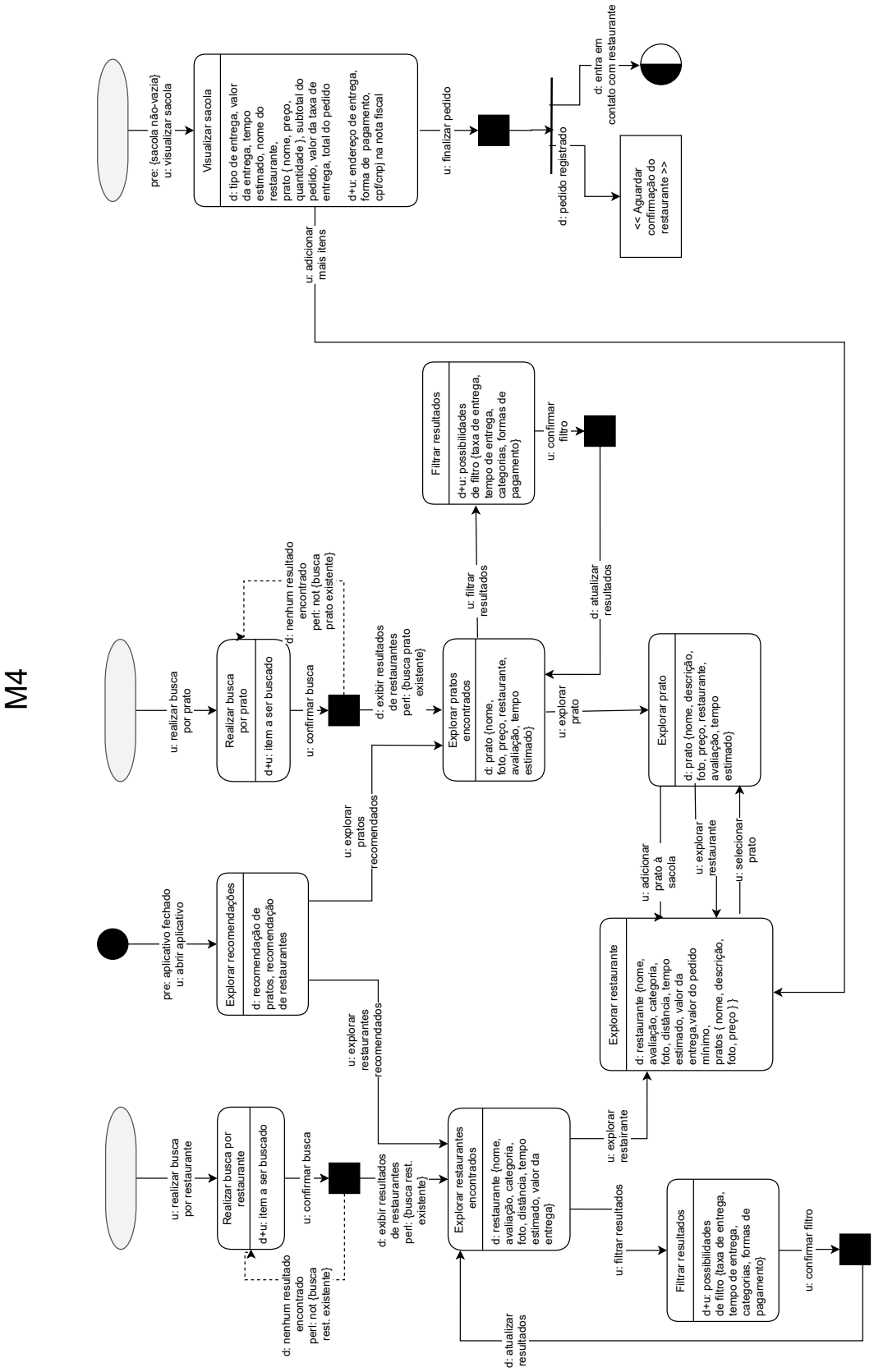

Figure C.4: M4

D
Study 2 - Alignments: MoLIC–Personas and MoLIC–Values

Marque o quanto cada diagrama MoLIC é apropriado para cada persona:

M ___

Janaína

Nada apropriado - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Muito apropriado

Sofia

Nada apropriado - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Muito apropriado

M ___

Janaína

Nada apropriado - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Muito apropriado

Sofia

Nada apropriado - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Muito apropriado

Marque o quanto cada diagrama MoLIC atende aos seguintes fatores:

M ___

Preço

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Familiaridade com o restaurante

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Higiene

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tipo de embalagem usada

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tempo de entrega

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tipo de comida

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Segurança dos dados

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

M ___

Preço

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Familiaridade com o restaurante

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Higiene

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tipo de embalagem usada

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tempo de entrega

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Tipo de comida

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

Segurança dos dados

Não atende - 0 ◯ ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 - Atende muito

E
Study 2 - Remarks made by participants about M1

Table E.1: Qualitative remarks made by participants about M1, per task

	T1	T2	T3	T4
Time	P2	P1, P2, P6, P14		
Price/discounts		P1		
Cost-benefit		P2		
Convenience	P2	P1, P14		
Flexibility				
Trust				
Exploratory behavior				
Comfort zone				
Familiarity				
Ratings	P6	P1, P6, P14		
Food quality		P1, P6, P14		
Distance		P1, P2		
Delivery fee		P1		
Food type		P1, P2		P9, P14
Satisfaction		P1, P14		
Vegan/vegetarian	P1, P5, P9, P13, P14	P1, P2, P5, P6, P9, P13, P14		
Other food restrictions	P5, P6, P9, P13, P14	P1, P2, P5, P9, P13, P14		
Sustainability	P5, P14	P1, P2, P5, P6, P9, P13, P14		
Packaging	P5	P1, P13, P14		P9
Health	P14	P1, P2, P9, P14		
Empathy				
Charity				
Social concerns				
Customization				

F
Study 2 - Remarks made by participants about M2

Table F.1: Qualitative remarks made by participants about M2, per task

	T1	T2	T3	T4
Time	P3, P8	P8		
Price/discounts	P3, P8, P10			
Cost-benefit				
Convenience	P3, P7, P8, P10, P11, P12	P8, P10		
Flexibility		P12		
Trust		P8		
Exploratory behavior	P8	P8		
Comfort zone	P7	P7		
Familiarity	P7			
Ratings				P12
Food quality	P11			
Distance				
Delivery fee				
Food type			P10	P7, P12
Satisfaction	P8, P11			
Vegan/vegetarian	P4	P3, P4, P10	P12	P8, P11
Other food restrictions	P4	P4	P12	P11
Sustainability	P10	P10		P11
Packaging	P4, P10		P12	P7, P8, P11
Health				P11
Empathy	P10	P4, P10		
Charity	P4, P10	P3, P4, P10	P12	P7, P8, P11
Social concerns	P10	P3, P4, P10	P12	P11
Customization	P4, P10			P11

G
Study 2 - Remarks made by participants about M3

Table G.1: Qualitative remarks made by participants about M3, per task

	T1	T2	T3	T4
Time	P3, P5, P13	P1, P5		
Price/discounts	P1, P3, P5, P7, P9, P11	P1, P7		
Cost-benefit				
Convenience	P3, P5, P9, P13	P1		
Flexibility				
Trust				
Exploratory behavior	P1, P7, P9	P7, P9		
Comfort zone				
Familiarity	P1			
Ratings		P1, P9		
Food quality		P1		
Distance	P5	P1		
Delivery fee		P1		
Food type			P1	P9
Satisfaction	P5	P1		
Vegan/vegetarian				
Other food restrictions				
Sustainability				
Packaging				
Health				
Empathy				
Charity				
Social concerns				
Customization				

H
Study 2 - Remarks made by participants about M4

Table H.1: Qualitative remarks made by participants about M4, per task

	T1	T2	T3	T4
Time	P2, P8, P10, P12	P2, P6, P8, P14		
Price/discounts	P8			
Cost-benefit		P2		
Convenience	P2, P8, P10, P12, P14	P8, P14		
Flexibility		P12		
Trust		P8		
Exploratory behavior	P2, P4, P8	P8		
Comfort zone				
Familiarity				
Ratings	P4	P6, P14		P12
Food quality		P6, P14		
Distance		P2		
Delivery fee				
Food type				
Satisfaction	P8	P14		
Vegan/vegetarian				
Other food restrictions				
Sustainability				
Packaging				
Health				
Empathy				
Charity				
Social concerns				
Customization				