Part I

Preference Representation

In this part, we address the problem of representing natural preference statements provided by users. The main goal of capturing and representing (user) preferences is to use them as input for a reasoning process that makes decisions on behalf of users. Examples of questions to be answered taking into account preferences are: (i) given a set of options, which is the most preferred? or (ii) how options are ranked according to user preferences? As preference representation models are used as input for algorithms that answer these questions, representation and algorithms are tightly coupled, because little can be done with preferences without having the ability to reason about them. On the other hand, considering that users have to state their preferences, there is no point in defining a preference representation model, which can be reasoned about, but users are not able to express preferences with constructions provided by this representation.

In fact, this issue was pointed out by Domshlak, who defines it as a paradoxical deadlock situation suggesting the "chicken-and-egg" metaphor. "*On the one hand, it is only natural to assume that reasoning about user's preference expressions is useful in many applicative domains (e.g., in online catalog systems). On the other hand, to our knowledge, no application these days allows its users to express any but trivial (e.g., "bag-of-word") preference expressions. It seems that the real-world players wait for the research community to come up with a concrete suggestion on how natural-language style preference expressions should be treated, while the research community waits for the real-world to provide it with the data essential to make the former decision. It is clear that this deadlock situation should somehow be resolved, and we believe that now this should be a primary goal for both sides.*" (Domshlak 2008)

In this context, our focus here is to understand how people explicitly express their preferences (without the aid of elicitation mechanisms), without being concerned with how to reason about stated preferences and make decisions based on them. Our goal is to define a preference representation model that serves as a reference for what should be ideally used as input for a preference reasoning algorithm, i.e. we aim at providing the essential data mentioned by Domshlak. We first present a study of how humans express preferences about a particular domain, from which we extracted patterns and expressions used by people to express their preferences (Chapter 2). Based on the analysis of preferences of our study, we propose a preference metamodel (Chapter 3). Finally, we present research work on preference representation models in different research areas from computer science (e.g. artificial intelligence and databases) and analyse their expressivity by identifying which kind of statements they are able to represent explicitly, which comprise only a subset of the statements identified in our study (Chapter 4).

2
Understanding User Ability to Express Preferences

Given that our approach requires users to express their preferences in a high-level language, we performed an exploratory study, presented in this chapter, to evaluate the feasibility of this requirement, and also to investigate how people express their preferences about a domain, including the common expressions they use. This study also evaluates the impact of experiencing a concrete decision-making situation and of the knowledge about the domain, while people state their preferences. We thus focus on answering two main research questions: (i) *are users able to express their preferences in such a way that a domain specialist is able to make an adequate choice in this domain on their behalf?* and (ii) *do users need to be exposed to a concrete decision-making situation, i.e. being aware of the available options, to be able to express their preferences about a familiar domain?* Other issues are also analysed, such as which kinds of changes users make after being exposed to a decision-making situation; and how the domain knowledge or other relevant aspects (age, gender, etc.) impact the users' expression of their preferences. If we conclude that the preference specifications given prior to a decision-making situation are not enough for making a decision on behalf of users — question (i) — it is essential to identify which kind of support can be provided for each user category in order for users to better express their preferences, but still without having to go through the entire decision process. Our study allowed us to identify kinds of support users need to better express their preferences and relevant concepts that should be part of an end-user preference language, which can be used by users to express their preferences in a way similar to natural language, or to correct and refine preferences initially acquired implicitly.

Our study consists of applying a questionnaire for participants of different profiles (knowledge about the target domain, age, working area, gender, and so on) to collect preference specifications expressed in natural language before and after experiencing a concrete decision-making situation. Later a domain specialist uses the initial specification to make recommendations according to each specification. Different measures are extracted from the collected data, both qualitative (e.g. type of preference specification and type of change) and quantitative (e.g. the amount of time and number of steps to make a decision). The domain chosen for our study is

Definition element	Our experiment goal
Motivation	To understand how people express their preferences,
Purpose	characterise and evaluate
Object	preference specifications
Perspective	from a perspective of the researcher
Domain:user	as people with different knowledge about a domain express their preferences
Scope	in the context of the researcher's social network.

Table 2.1: Goal Definition (GQM template).

laptop purchasing. This choice was made because choosing laptops represents the kind of task we are addressing in our work: users are aware of a set of preferences over laptops, as buying laptops is a task that is potentially performed repetitively (every x years), but each time it is different as the available options and features evolve over time. In addition, users have different levels of knowledge about this domain and we had a domain specialist available to participate in the experiment. The main goal of this study is to give foundation to our work on automated decision making.

The remainder of the chapter is organised as follows. Sections 2.1 and 2.2 describe the design and results of our study, respectively. Section 2.3 discusses the results, followed by Section 2.4, which concludes.

2.1 Study Description

In this section we detail the design of our exploratory study, as well other relevant information, including the research questions we aim to answer and the participants involved.

In order to design our study, we have followed the framework proposed by Basili et al. (Basili et al. 1986), which provides guidelines to elaborate experimental studies in Software Engineering (SE). The first phase of the framework is the definition of the experiment adopting the goal-question-metric (GQM) template (Basili and Rombach 1988), which establishes experiment goals that are used for defining research questions associated with it. Then metrics are defined or selected for answering those questions. Following this template, the experiment goal is presented in Table 2.1. After that, the phases of planning, operation and interpretation are executed. Both (Basili et al. 1986) and (Basili and Rombach 1988) provide guidance for performing experimental studies in the context of SE, but they are generic enough to be used in our study. Their adoption was due to our previous experiences with SE studies.

Our study consisted of applying a web-based questionnaire for profiling the participants, capturing their preferences on a domain before a decision-making situation, engaging participants in concrete decision making in that domain, and then asking them to review their previous preference specification. Next, a domain expert made recommendations for participants based on their initial preference specifications and the same available options, and all the collected data was then analysed. The domain chosen for our study is laptops. This decision was made due to the availability of domain experts, and also because this domain illustrates a scenario in which participants might have experienced a similar decision-making situation, but always with different available options, as laptops evolve over time, and new features are introduced. More details of the study procedure are given in Section 2.1.2.

2.1.1 Research Questions

The main goal of this study is to evaluate how users would typically express their preferences about a domain, and how useful the provided preference specification is to make a decision on their behalf. In addition, we aim to investigate the impact of experiencing a concrete decision-making situation on this specification. This evaluation was performed in different directions, which are associated with seven research questions addressed in the study, as presented in Table 2.2(a).

With these research questions, we aim to acquire a deeper knowledge on user preference expression. This information enables us to make statements about users, and it is helpful and necessary for developing approaches in the context of preference-based decision making. In particular for our approach, it allows us to verify whether it makes sense to provide an end-user language for users to express or adjust their preferences so they can delegate tasks to systems provided with automated decision making. If users are unable to specify their preferences in natural language in such a way that a (human) expert in that domain is able to make an appropriate choice on their behalf, as it is the case in our study, it is unlikely that it will work with a restricted language and software systems. This issue is addressed by *RQ1*. In addition, we also investigate whether users need to experience a concrete decision-making situation, i.e. they need to know the available options, in order to adequately express their preferences (*RQ2*).

Moreover, we also address other issues with the elaboration and execution of our study, which can help in different aspects of the development of preference-based approaches. With *RQ3*, we aim to identify which problems (wrong, missing or outdated preferences) occur when users specify their

RQ1. Are users able to express their preferences about a familiar domain in such a way that a (human) domain specialist is able to make an adequate choice in this domain on their behalf?	**EA1.** Comparison between laptops selected by participants and the ones recommended by the domain specialist based on their specification.
RQ2. Do users need to be exposed to a concrete decision-making situation to be able to express their preferences about a familiar domain?	**EA2.** Analysis of the differences between the initial preference specification and the reviewed version of it.
RQ3. Which type(s) of preferences users usually forget or incorrectly specify before being exposed to a concrete decision-making situation?	**EA3.** Analysis of the most common types of preferences that appeared only in the preferences review.
RQ4. How different are specifications provided by users with a high degree of knowledge about a domain from the ones provided by those with a low degree of knowledge?	**EA4.** Comparison between the preference specifications provided by participants with high degree of knowledge about the domain and by those with low degree of knowledge about it.
RQ5. Which user profiles take less time to express their preferences?	**EA5.** Comparison of how long participants classified in different categories (domain knowledge, gender, ...) take to specify their preferences.
RQ6. When users make a choice, which ones select fewer options from among the offered ones? In other words, which user profiles are more confident in which is the right choice for them?	**EA6.** Comparison of how many laptop options were chosen by participants in different categories (domain knowledge, gender, ...).
RQ7. Which user profiles take less steps (filtering, comparing, analysing, ...) in the decision-making process (choosing among available options)?	**EA7.** Comparison of how many steps (filtering, looking details, comparing,...) participants in different categories (domain knowledge, gender, ...) took to define their laptop options.

(a) Research Questions. (b) Evaluation Approaches.

Table 2.2: Research questions and their evaluation approach.

preferences, so that we can provide mechanisms to prevent them. With *RQ4*, we investigate whether and how users with different knowledge about the domain express their preferences in different forms; if so, languages with different vocabularies might be needed for different user profiles. As the effort that users spend to perform a task is directly related to how much they are willing to do it, they might not want to do it. Therefore, in order to evaluate the feasibility of expecting users to provide preferences in a high-level language, we evaluate in *RQ5* whether users (or some of them) take too long to express their preferences. Confidence on the choices made is relevant to investigate (*RQ6*), as it can indicate that different options may satisfy users' needs or that trade-off situations were not resolved, and different approaches may be adopted by automated decision making systems to include users in the decision making process, according to their confidence on the decision. Finally, *RQ7* helps to understand the user decision-making process, and how it differs when the user has deep knowledge about the domain. This is also related to the elaboration of approaches that make decisions on behalf of users.

2.1.2 Procedure

The study we planned to answer our research questions is mainly based on a web-based questionnaire applied to a wide spectrum of users (see next section for details). The domain selected for performing our study is shopping for products; in particular, we chose the laptop as the target product. This decision was made due to the reasons introduced before.

In a nutshell, the idea of the questionnaire is to first ask users to specify their preferences for someone who is going to buy a laptop for them. Later, they are asked to navigate on a laptop catalog and select from one to five laptops. Finally, we give users a chance to modify their preference specification.

The applied questionnaire, which can be seen in Appendix A, consists of four parts, each of which is explained next.

- *User Information Data.* The questionnaire is anonymous, but we collect relevant information related to the study from the participant: (i) age; (ii) location (city and country); (iii) working/studying field; (iv) how many laptops the participant has already had (current one included); (v) from these, how many were chosen by the participant herself; and (vi) how she rates her knowledge about the domain. These last three items are used to evaluate the participants' knowledge about the domain.
- *Preference Specification.* The study participant is requested to imagine a situation in which she is going to ask someone to buy a laptop for her. Therefore, she is requested to specify all her preferences and restrictions. An example in the flight domain is provided in order to give some instructions for participants, as they are not assisted while answering the questionnaire. An example in this domain was adopted because we did not want to influence the participants by providing an example in the same domain of the study, and also because the process of choosing seats has similarities with the process of choosing laptops: available seat locations vary each time a decision is made. Besides storing the provided preference specification, we logged the current state of the specification every 15 seconds and the time the participant took in this part of the study.
- *Choosing Product.* Next, the participant is requested to analyse a set of different computers and say which one she would have bought. We ask her to rank her favourite ones, up to five laptops. We used the Best Buy[1]

[1] http://www.bestbuy.com/

catalog, which had 144 laptops by the time we imported it (at the same day in which the survey was released). We recorded each step (comparing, filtering, detailing, ...) the participant performed, as well as the time taken for choosing the laptops.

– *Preference Specification (review).* Finally, after analysing the available computers, the participant is given a chance to review her preferences and modify them, in case she realised that something was missing or wrong in her specification. We have notified participants in the third part that they would have this reviewing chance. We also asked the participant's comments on what changed on her specification. The additional logs are the same as in the second part of the questionnaire.

After collecting all the data, a domain expert was involved in the study. The domain expert's responsibility was to analyse the first version of the preference specifications provided by the participants, and to rank up to five laptops he would have recommended for each one. We are aware that involving more than one domain expert to make recommendations would significantly improve the results of our study, but we did not find other experts willing to participate in it.

The study participants and the domain expert were allowed to choose up to five laptops, which is a limit we established. We chose this number because we wanted to provide flexibility for participants and for the expert to choose more than one laptop, and we assumed that five options were enough. In order to confirm that five is a good number as well as to evaluate our web interface, we executed a pilot study with few individuals, and they approved both the interface and the limit of five options. In addition, as it can be seen in the results, several participants selected less than five laptops. With regard to the domain expert recommendations, we asked him whether he wanted to recommend more than five options and he answered that five was a good number.

Based on the questionnaires and the recommendations made by the domain expert, we analysed the data according to two main aspects, related to the research questions 1 and 2: (i) were the participants able to express their preferences in such a way the domain expert could make adequate recommendations for them? and (ii) did the participants change their preference specifications after experiencing the process of choosing a computer? Furthermore, we have also analysed other relevant aspects in order to answer the additional research questions, from 3 to 7. In Table 2.2(b), we detail how we analysed the survey data to answer each research question.

The evaluation approach presented in Table 2.2(b) shows we have performed mainly a subjective but also an objective analysis of the data to answer all our research questions. Table 2.3 summarises all collected data, both qualitative and

Subjective	Objective
• Similarity score between chosen laptops and those recommended by the domain expert • Types of preference changes • Types of preference specifications • Characteristics of preference specifications	• Time taken to specify preferences • Number of chosen options • Steps taken to choose options • Number of participants that changed their specifications

Table 2.3: Qualitative and quantitative data collected.

quantitative, from our study. Some of the measures, e.g. similarity score and types of preference changes, are described later, when they are used to analyse results.

2.1.3 Participants

Our study involved a total of 192 participants, who answered our questionnaire, and one domain expert, who indicated laptops for each participant according to their initial preference specification. Due to the effort needed to analyse a large number of questionnaires, only one domain expert agreed to participate in our study.

The questionnaire was available online from May 20 to July 13, 2010 (almost two months). For selecting the participants, we used convenience sampling, based on the social network of the researchers involved in this study. First, invitations for participating in the study were sent by e-mail and people were asked to forward the invitation for other people in a snowball approach. In addition, a call for participation in the study was published in different Orkut[2] communities.

As result, we collected a database with 451 surveys that were initiated, from which 192 were completed (42.6%) — incomplete surveys were discarded. As the researcher that performed this study is Brazilian, most of the participants are from this country (86.98%), and the remaining ones (13.02%) are from four other countries. The same situation happens with the working area (63.54% participants work with a background in informatics-related areas). In our analysis, we did not detail other working areas (despite having this information available in our database) because our focus was to identify participants with a higher knowledge about our study domain (laptop purchasing). The description of the demographic characteristics of our study participants is detailed in Table 2.4.

The domain expert that was involved in our study has an M.Sc. degree in Computer Science. Moreover, his work involves giving technical support to the Software Engineering Laboratory at PUC-Rio as well as specifying and

[2] `http://www.orkut.com`

Working Area	Informatics 122 (63.54%)	Non-informatics 70 (36.41%)	**Gender**	Male 134 (69.79%)	Female 58 (30.21%)
Domain Knowledge	No Knowledge 5 (2.60%)	Beginner 16 (8.33%)	Intermediate 40 (20.83%)	Advanced 83 (43.23%)	Expert 48 (25.00%)
Country	Brazil 167 (86.98%)	Germany 10 (5.21%)	Canada 10 (5.21%)	United States 4 (2.08%)	Peru 1 (0.52%)
Age	16-25 years 60 (31.25%)	26-35 years 83 (43.23%)	36-45 years 21 (10.94%)	>45 years 28 (14.58%)	

Table 2.4: Demographic Characteristics of Participants.

recommending new computers and laptops for the laboratory and its individual members. Therefore this expert is used to listening to clients specifying their preferences and to recommending computers and laptops for them.

2.2 Results and Analysis

In this section we provide the results we collected from the execution of our study as well as interpretations for those results. We have made a qualitative analysis of the preference specifications (initial and revised versions) given by the study participants and a quantitative analysis of part of the collected data, such as time taken to answer each part of the questionnaire.

The first analysis that we made was how to measure the participants' knowledge about laptops. The fields (iii) to (vi) in the *User Information Data* part of the questionnaire were used for that. Based on this data, we make the following observations.

– Participants that work in the computer science area have at least an INTERMEDIATE[3] level of domain knowledge. In other fields, participants are mostly INTERMEDIATE.

– Most of the participants who have had several laptops have at least an ADVANCED knowledge; the more laptops participants have had, the higher their knowledge.

– Almost all of the participants chose their laptops; only the ones who had several laptops had some laptops chosen for them (possibly because they get laptops from their companies).

– Not having had a laptop does not indicate a low knowledge about the domain — some participants chose not to have a laptop.

The relationship between the other fields and the domain knowledge provided by the participants presented the behaviour we expected: participants with a background in informatics-related areas or those that had many laptops (and chose

[3]Values for evaluating the domain knowledge: NO_KNOWLEDGE, BEGINNER, INTERMEDIATE, ADVANCED, EXPERT.

them) rated themselves with expertise on laptops. Therefore, when analysing the collected data, we adopted the domain knowledge that the participants themselves provided as a resource to determine their knowledge about the domain.

RQ1. Are users able to express their preferences about a familiar domain in such a way that a (human) domain specialist is able to make an adequate choice in this domain on their behalf? Based on the initial preference specifications of all participants, and the domain specialist recommendation for each of them, we have compared the laptops recommended and the ones the participants chose. The goal was to investigate the users' ability to express their preferences and how their knowledge about the domain influences this ability. Some participants, instead of providing a specification, stated that they would never delegate this task to a person, or provided templates using variables for referring to attribute values of the laptops, e.g. *"I would like laptops with processor X."* Without a specification, the domain expert is not able to make a recommendation, therefore nine of the surveys were discarded for this research question. From 183 surveys, 53 (28.96%) had at least one of the specialist's recommendations that matched at least one of the participants' choices. Besides analysing exact recommendation matches, we further calculated the similarity between recommendations and participants' choices. As we have up to five laptop choices for both the domain expert and participants, we did not calculate this similarity by simply comparing one selection with another, such as calculating the mean square error of individual laptop features. We thus have elaborated a function — shown below — to calculate this similarity score (SS), which takes into account the positions matched to calculate a weighted average.

$$SS = \frac{\sum_{i=0}^{size(CL)}(5 - i) * \left(\frac{\sum_{j=0}^{size(SR)}(5-|i-j|)*(sim(CL[i],SR[j]))}{\sum_{j=0}^{size(SR)} 5-|i-j|} \right)}{\sum_{i=0}^{size(CL)} 5 - i} \qquad (2\text{-}1)$$

where CL is the chosen laptops (by the participant), SR is the specialist recommendation (for the participant), $size(v)$ returns the size of a vector v and $sim(x, y)$ is the function that calculates the similarity between two laptops. If they are equal, its value is 100, otherwise it is the average of each feature compared. If the feature has a numeric domain, the feature comparison is $|featureValue1 - featureValue2| / (upper\ bound - lower\ bound)$, where the domain boundaries are given by the highest and lowest values for the feature considering all laptops. Otherwise, the feature comparison is 100 for equal values, 50 for unspecified values, and 0 for different values. Table 2.5 presents the values found in our study, which ranged from 47.87 to 100.00. The column *matches* is the number of surveys in which at least one of the laptops matched, and the columns *SS(M)* and

	Matches	SS(M)	SS(Median)	SS(SD)
Domain Knowledge				
NO_KNOWLEDGE	3 (60.00%)	68.58	63.76	18.45
BEGINNER	2 (13.33%)	59.14	58.90	4.37
INTERMEDIATE	9 (23.08%)	60.93	59.39	6.31
ADVANCED	27 (33.33%)	62.82	60.30	8.38
EXPERT	12 (27.91%)	62.73	58.96	9.41
Gender				
FEMALE	17 (29.82%)	61.65	59.64	7.62
MALE	36 (28.57%)	62.52	59.81	8.79
Age				
16-25 years	14 (24.14%)	61.51	59.16	8.74
26-35 years	27 (35.06%)	63.28	60.27	9.10
36-45 years	8 (38.10%)	63.70	62.05	7.99
>45 years	4 (14.81%)	59.78	59.41	5.17
Total	**53 (28.96%)**	**62.25**	**59.76**	**8.43**

Table 2.5: Domain Specialist Recommendation — Matches per Group of Participants.

SS(SD) are the average and standard deviation of the similarity score, respectively. In Table 2.5, we show not only these numbers of the SS for all participants, but also for different group categories.

Considering the SS values, it can be seen that the domain expert was not able to match laptops very closely to participant choices — only 19 participants had $SS > 70$, 11 had $SS > 80$, 4 had $SS > 90$, and 28.96% matches — however he managed to recommend laptops that have at least half of the characteristics selected by participants. This result indicates that specifications explicitly provided by users are valuable sources for providing recommendations for them, as in some cases they allow the specialist to make adequate decisions. However, for making decisions on their behalf, users must be instructed on how to provide better specifications or additional support should be provided for improving their specifications. Research question 4 indicates some problems identified in the preference specifications.

Table 2.6 presents the number of matches according to each rank position of the laptops chosen by participants. For some participants, more than one position matched. It can be seen that the number of matches is higher in the first positions. This result corresponds to the quality of the specification: when the specification provides good details of what users want, it is more likely that the exact laptop they want is matched.

Position Matched	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}
#Matches	30	17	11	8	2

Table 2.6: Domain Specialist Recommendation — Position Matched.

Table 2.5 also shows that the number of matches is higher for participants with a higher knowledge about the domain (ADVANCED and EXPERT). This can

Domain Knowledge	Who Changed (%)	#Changes (M)
NO_KNOWLEDGE	0 of 5 (0.00%)	0.00
BEGINNER	4 of 16 (25.00%)	2.50
INTERMEDIATE	14 of 40 (35.00%)	2.29
ADVANCED	28 of 83 (33.73%)	3.04
EXPERT	16 of 48 (33.33%)	2.13
Gender	Who Changed (%)	#Changes (M)
MALE	47 of 134 (35.07%)	2.53
FEMALE	15 of 58 (25.86%)	2.80
Age	Who Changed (%)	#Changes (M)
16-25 years	27 of 60 (45.00%)	2.07
26-35 years	25 of 83 (30.12%)	3.12
36-45 years	6 of 21 (28.57%)	2.33
>45 years	4 of 28 (14.29%)	3.25

Table 2.7: Preference Changes.

also be seen in the similarity score. Nevertheless, the highest number of matches (in percentage) were in the group of NO_KNOWLEDGE participants. We observed that these specifications, even though they do not contain specific details of the laptop, provide key information about the purpose for which the laptop will be used. But it is important to highlight that we are not aware of which criteria the participants used to choose the laptops, as they do not have knowledge about the domain.

In order to test whether the difference among the matches for the groups with different domain knowledge is statistically significant, we used the Kruskal-Wallis test. The recommendations *did not differ* significantly across the five levels of domain knowledge, $H(4) = 3.755$, $p = 0.4402$.

RQ2. Do users need to be exposed to a concrete decision-making situation to be able to express their preferences about a familiar domain? From the 192 participants, only 62 (32.29%) modified their preference specification after experiencing choosing laptops and navigating through the catalog. This result shows that even though the preference construction for available laptops, i.e. establishing an order for the available options, was made when participants had to choose one (or some) of the laptops, they did not change their preferences for individual attributes, and any method that they may used for resolving trade-off among attributes was not reported as a preference. Table 2.7 shows the participants that changed their preferences according to the domain knowledge, gender and age. In addition, it presents the average number of changes (we explain how we counted it in the next research question).

From these 62 participants, no one had NO_KNOWLEDGE about the domain. Even after searching laptops and seeing their features, NO_KNOWLEDGE participants were unable to describe preferences in terms of the laptop features — they did not know (and maybe did not want to know) what these features mean. A

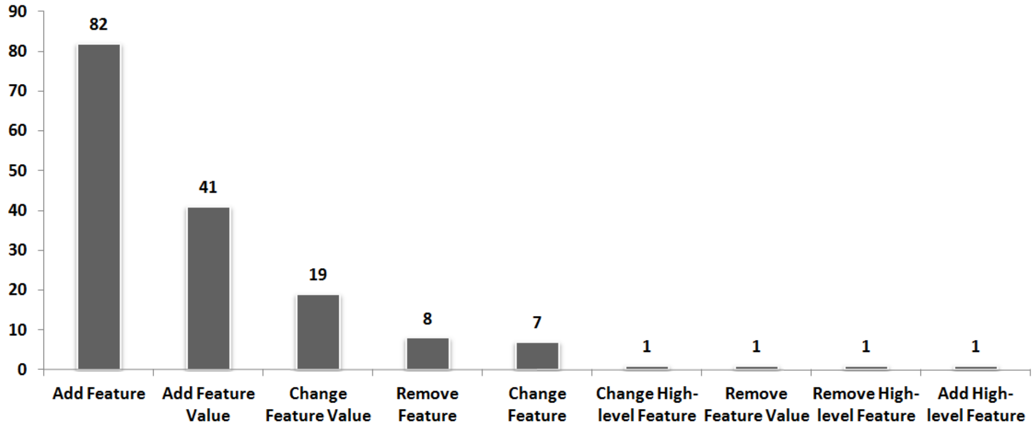


Figure 2.1: Nature of Preference Changes.

few BEGINNERS changed their specification, but they added high-level features such as "modern design" and "installed software," and particular features learned from the catalog did not influence their specification. Approximately one third of the three remaining categories changed their specification. They understand the domain (some of them better), but not necessarily know the latest news (this was the main reason for the changes made by EXPERTS). When they see new and updated features, or features they forgot to mention, they provide further details on their specification.

Analysing changes and ages, the older the participant is, the fewer changes she made. Older people provided less *detailed* specifications (see research question 4), but still did not change them after going through the process of decision-making. However, when they changed their specification, they made more changes.

RQ3. Which type(s) of preferences users usually forget or incorrectly specify before being exposed to a concrete decision-making situation? We analysed all specifications that changed when they were revised, and classified each change with a target and a type. There are three kinds of types: *add*, *remove*, or *change*. Also, there are three kinds of targets: (i) *Feature*: it describes a characteristic of the laptop, e.g. "HDMI;" (ii) *Feature value*: it describes the value of a feature, e.g. "Processor i5" changed to "Processor i5 or i7;" (iii) *High-level Feature*: it describes a high-level characteristic of the laptop, e.g. "Mobility." When a participant added a feature *and* its value in the preferences review, it was considered as a feature, because the feature would not make sense without a value. But if the participant only added a value to an existing feature, it was considered as add feature value. Figure 2.1 shows the occurrence of preference changes according to their nature (target and type).

As it can be clearly seen, the three most common types of preference change

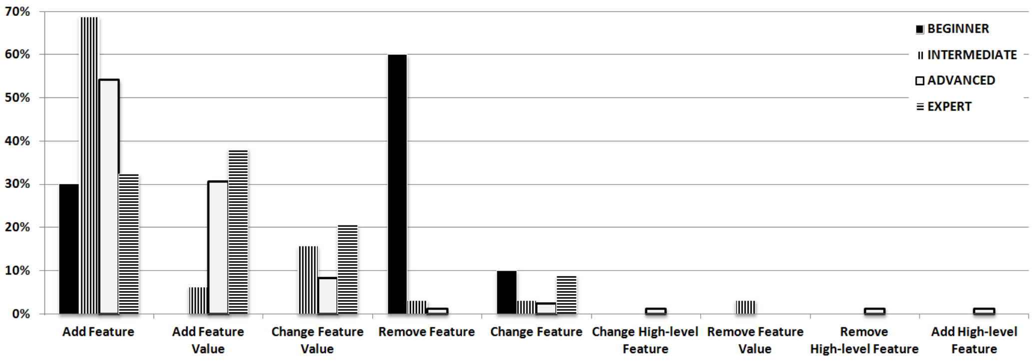


Figure 2.2: Preference Changes x Domain Knowledge (Percentage).

that participants made in the preference review are: (a) Add Feature (50%); (b) Add Feature Value (25%); and (c) Change Feature Value (12%). What happened was that users forgot to specify some characteristics that are important for them, or there is a new characteristic that they did not know about. At the moment they saw them in the laptop catalog, they remembered to specify them.

Moreover, some of the users were not aware of the current average or top values (price, processor, etc.), and as they know this by searching an up-to-date catalog, they realise that the value is different from what they thought (it is mainly related to feature values). However, some participants specified feature values in terms of relative values ("second best value"), instead of absolute ones ("4GB"). Using this kind of specification makes the preference specification reusable in different occasions.

Figure 2.2 presents how preference changes occurred distributed across the different domain knowledge categories, where it can be seen the three most common types of preference change presented above is the same for almost all categories. The only exception is in the BEGINNER category, in which 60% of the changes are of the type remove feature. However, it happened because a single participant changed the way she provided her specification, and therefore she removed the previously provided features and added a different kind of information (provided a specific laptop model).

RQ4. How different are specifications provided by users with a high degree of knowledge about a domain from those provided by users with a low degree of knowledge? The goal of this research question is to investigate how users with different knowledge about the domain express their preferences. As users that do not know much about laptops are not aware of their features, they tend to use an alternative vocabulary in comparison with domain experts.

We analysed each preference specification and classified it in four different types, which take into account only the laptop specific features. In addition, we have

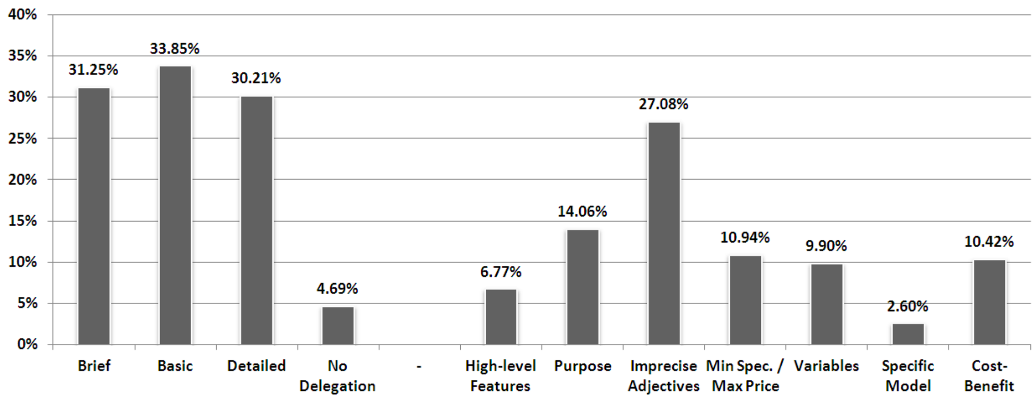


Figure 2.3: Preference Specification Analysis.

identified particular characteristics and common patterns. Figures 2.3, 2.4 and 2.5 present the charts that show the data collected (percentage) from the preference specifications from our study. Figure 2.3 shows the results related to the whole group of participants, while Figures 2.4 and 2.5 show two perspectives from the specifications, their type and their characteristics, classified according to three different categorisations: (i) domain knowledge — Figures 2.4(a) and 2.5(a); (ii) gender — Figures 2.4(b) and 2.5(b); and (iii) age — Figures 2.4(c) and 2.5(c). The four types of preference specification are presented below.

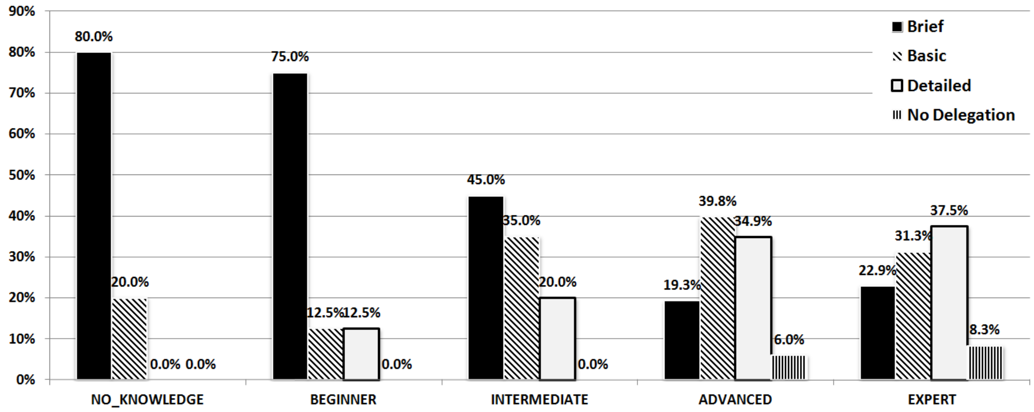
- **Basic** specifications mention characteristics for features that are part of every laptop (processor, RAM memory, hard drive, screen size). Characteristics can be specific values, or adjectives, such as “good” and “big.” Example:
  Brand HP or Dell
  Processor Core 2 Duo or better
  RAM Memory 4 GB or more
  HD 300 GB or more
  Monitor of 15” or more
- **Brief** specifications do not cover laptop basic features (they mention none or few of them). Usually other kind of specification is provided, such as for what the participant will use the laptop. Example:
  I want one that is light
  - Screen of 14 inches
  - fast
  *- beautiful^^*
- **Detailed** specifications give further details about laptops than the basic specifications, i.e. they are more specific, tending to narrow the laptops search space. We added to this category brief and basic specifications of *Apple* laptops, namely **Brief but Enough**, because participants who want laptops

of this manufacturer, by describing only a few features, already indicate a specific laptop. Example of a detailed specification:
1 - The laptop must be of a traditional brand, preferentially Sony, Dell or Apple.
2 - Hardware configuration must be up-to-date (processor, RAM memory, HD, video card and video output, etc.).
3 - The operating system must be Windows 7 or Leopard Snow.
4 - The screen should preferably have 13” (it can have 14” if the laptop is very good and the price is attractive).
5 - I don’t care about buying an “open box” if the computer is good and the price attractive.
6 - The keyboard must be comfortable.
7 - The looks and design of the laptop must be sober and of a good taste.
8 - At least 2 USB ports and there must be an integrated microphone and camera.
9 - The audio output for earphone must be at the side of the computer (it cannot be in front of it). It is important also that it has volume control on the keyboard or in the chassis.
10 - The laptop must have a competitive price.

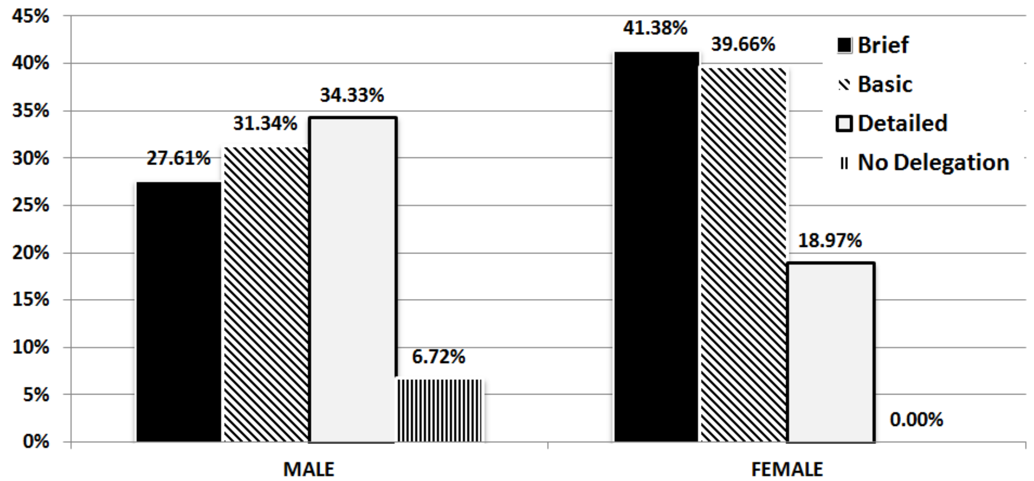
– **No Delegation.** Few (nine) participants did not provide a descriptive specification but informed the specific model they wanted, or stated that they would never delegate the decision for another person. Example:
Buy the Sony VPC F1190X, with the following configurations:
- Proc: Core i7 820QM 1.73GHz
- OS: Win7 64bits
- HD: 500GB 7200RPM
- Memory: 8GM DDR3 (1333MHZ)
- Drive Blu-Ray
- Monitor 16.4”
- VGA: GeForce GT 330M (1 GB of video)
- No additional software.

This categorisation emerged from the qualitative analysis of collected specifications, which is supported by principles of grounded theory (Glaser 1992), which is a systematic methodology for generating a “theory” from collected data. By observing the preference specifications, we identified patterns among them, and noticed that they could be classified in this way.

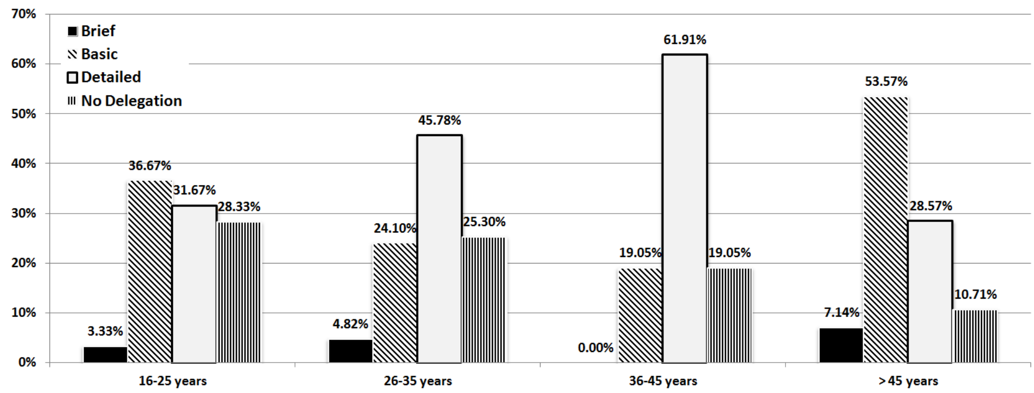
As stated previously, the presented categorisation has taken into account only how laptop specific features were described. We also observed other characteristics of the specifications, which are described next.



2.4(a): Domain Knowledge.

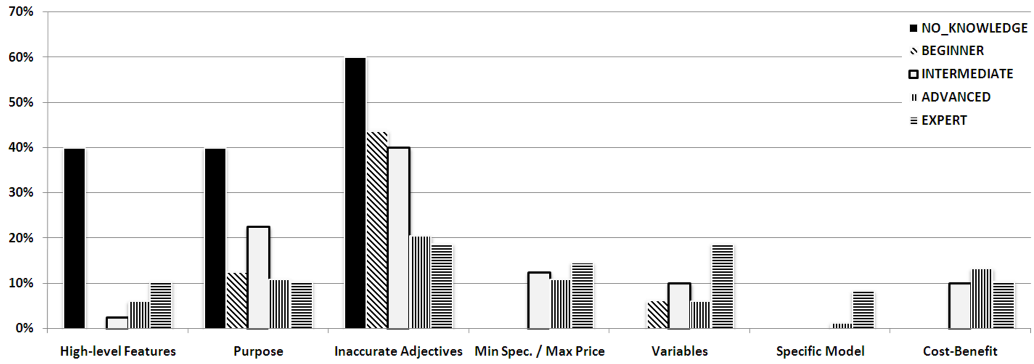


2.4(b): Gender.

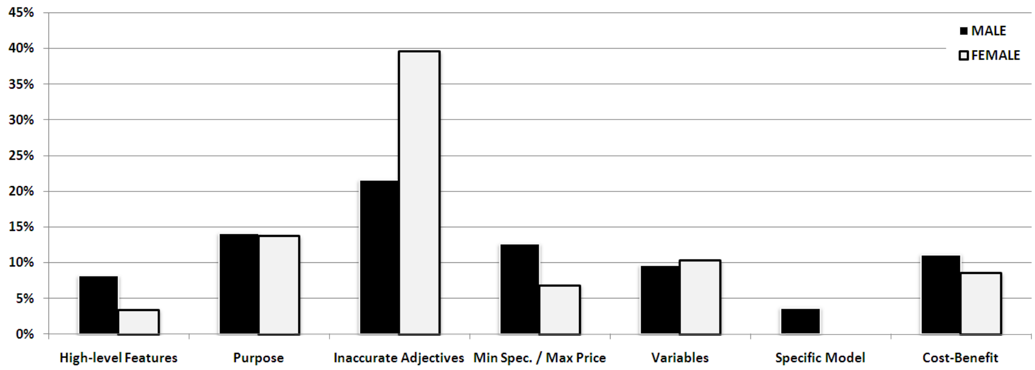


2.4(c): Age.

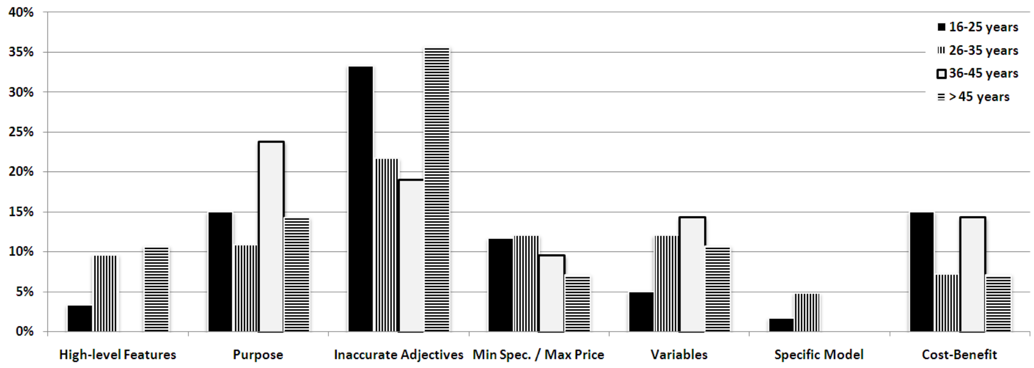
Figure 2.4: Preference Specification Types (percentage).



2.5(a): Domain Knowledge.



2.5(b): Gender.



2.5(c): Age.

Figure 2.5: Preference Specification Characteristics (percentage).

- Presence of **High-level Features**, which describe the consequences of having a value for a (set of) specific laptop feature, e.g. mobility, readability, performance. Example: *"Portability is an important factor — I'm not looking for a desktop replacement, but rather something that I can take with me when I travel."*
- Description of **Purpose** — specifications that contain the purpose for which the participant wants the laptop, for instance playing games. Example: *"I like playing not so hardware demanding games, however with good quality"* and *"My notebook is a desktop replacement."*
- Presence of **Imprecise Adjectives**, which are adjectives whose meanings depend on the point of view of the participant, e.g. "good," "modern design," "beautiful." Example: *"3. Not too heavy."*
- **Minimum Specification/Maximum Price** — pattern of specification that specifies a minimum specification for the laptop features and establishes a maximum price that the participant is willing to pay.
- Presence of **Variables**, which is when the participants used variables for feature values on their specification. Example: *"I would like laptops with processor X or Y with a screen size Z with optical drive K with hd of size L and memory of size A and video card F."*
- **Specific Model** — specifications that do not describe laptops but indicate the specific model the participant wishes. Example: *"Apple MacBook Air With 13.3 Display Aluminum."*
- **Cost-benefit** — participants that mentioned this characteristic on their specifications. Example: *"5. Look for the best cost vs. benefit among the laptops that fill the above specifications."*

We observed that participants with high degree of knowledge about the domain express themselves with fine-grained features (e.g. laptop-specific features) and participants with low degree of knowledge tend to refer to high-level features, as it can be seen in Figure 2.4(a), in which the level of detail of the preference specifications increases as the knowledge about the domain grows. Participants with a lower degree of knowledge specify their preferences without detailing too much the specific features of the laptop. They use high-level features to describe what they want and for what they need a laptop. Some of them mention that they would ask a friend who understands the domain for receiving a recommendation — these are not even interested in learning about the domain. On the other hand, participants with a higher degree of knowledge were much more specific, stating the exact values (or a range) for most of the laptop specific features. The level of precision in defining

	SS(**M**)	SS(**Median**)	SS(**SD**)
Brief	60.79	59.06	7.40
Basic	61.48	59.79	7.07
Detailed	64.14	61.73	9.71
No description	75.53	75.53	22.39
Total	**62.25**	**59.76**	**8.43**

Table 2.8: Preference Types vs. Domain Specialist Recommendation.

feature values decreases as the domain knowledge increases, but even EXPERTS use inaccurate adjectives (18.75%).

Even though this specification is supposed to instruct an individual to execute a task for the participant, there is a certain degree of autonomy — choosing the laptop. Some participants (6.02% ADVANCED and 8.33% EXPERT) did not provide a specification but gave the exact model they want. One of the participants stated *"I would never delegate such a decision to someone else."* This shows a group of people that do not trust other parties to decide on their behalf (at least for certain tasks). However, there are still other kinds of support that could be provided, such as making recommendations from which users could choose and make the final decision or checking prices in different stores, as stated by one of the participants: *"SEARCH other stores, before buying it."*

We have analysed the similarity score of the domain specialist recommendation used previously with respect to these preference types, in order to know how useful they were to make choices on the participants' behalf — the results are summarised in Table 2.8. We then used the Kruskal-Wallis test, performing the post-hoc tests of Nemenyi-Damico-Wolfe-Dunn. The results show that the similarity scores significantly differs across the different specification types, $H(3) = 7.882$, $p = 0.0485$. And the posthoc analysis shows us that the difference is due to the difference between the **Brief** and **Detailed** specifications (p-value= 0.0492).

We have also analysed the effect of age and gender on the preference specifications. However, the data obtained does not allow us to conclude significant difference among these different groups. An initial investigation indicates that differences of specifications provided by participants of different ages or genders are related to their levels of domain knowledge. Further investigations about it are outside the scope of this study.

RQ5. Which user profiles take less time to express their preferences? Table 2.9 shows how long (average, median and standard deviation) participants took for providing their initial preference specification, according to their domain knowledge, gender and age. We have observed different task times among participants with different domain knowledge.

		Specification Time		
		M	Median	SD
Domain Knowledge	NO_KNOWLEDGE	04:30	02:42	05:38
	BEGINNER	04:57	03:40	04:05
	INTERMEDIATE	06:35	05:12	05:11
	ADVANCED	06:45	05:49	05:28
	EXPERT	06:14	04:44	07:57
Gender	MALE	05:33	04:12	04:47
	FEMALE	06:44	05:21	06:27
Age	16-25	06:09	05:14	04:37
	26-35	06:32	05:17	05:53
	36-45	05:49	05:11	03:03
	>45	06:49	04:03	09:48

Table 2.9: Time Taken for Specifying Initial Preferences.

Participants with NO_KNOWLEDGE or BEGINNER domain knowledge took less time for building their specifications. One reason for this is that their specifications are smaller than the others'. Second, their specifications contain details about the purpose for which they need the laptop or high-level specifications, which are details that may be easier to remember. The participants who took longer specifying what they wanted were those with INTERMEDIATE or ADVANCED knowledge. Their specifications are more detailed, but they did not promptly remember what they wanted (we observed that in the specification logs). Sometimes they went backward and changed or added details to their specifications. Finally, EXPERT participants also provided detailed specifications, but as they are more familiar with the domain, their preferences have come easier to their mind.

RQ6. When users make a choice, which ones select fewer options from among the offered ones? In other words, which user profiles are more confident in which is the right choice for them? When users know exactly what they want — they are confident that there is one option that is best for them — they choose one option from the available laptops. Therefore, by observing the number of options chosen by participants, we can have an idea of their confidence while making the choice. So, regarding the number of laptops chosen by participants, we can observe that no group has an average or a median of less than three (see Table 2.10). It indicates that even when an individual knows the domain very well, there are different options that satisfy her needs. In addition, BEGINNER and INTERMEDIATE participants have a slightly higher average and median than the other categories of domain knowledge. Possibly, they do not care about minor details of the laptops, as ADVANCED and EXPERT participants do.

There is an existing framework (Chen and Pu 2010) that considers the *objective* accuracy as one of the criteria for evaluating recommender systems, which compares the final option found with the decision tool to the target (best) option that

		Options		
		M	Median	SD
Domain Knowledge	NO_KNOWLEDGE	3.40	3.00	1.67
	BEGINNER	3.88	5.00	1.54
	INTERMEDIATE	4.10	5.00	1.30
	ADVANCED	3.67	4.00	1.44
	EXPERT	3.21	3.00	1.53
Gender	MALE	3.74	4.50	1.49
	FEMALE	3.62	4.00	1.46
Age	16-25	3.60	4.00	1.55
	26-35	3.75	4.00	1.41
	36-45	3.76	4.00	1.37
	>45	3.43	3.50	1.55

Table 2.10: Number of Chosen Laptops.

users find after reviewing all available options in an offline setting. However, as our participants did not chose only one laptop, it might lead to the conclusion that such “best option” does not exist. In the field of marketing, it is more common to talk about “client satisfaction,” which is more related to the *perceived* accuracy criteria of the framework.

RQ7. Which user profiles take less steps (filtering, comparing, analysing, ...) in the process of decision-making (choosing among available options)? Besides storing the laptops chosen by participants, we have also logged their actions each time they executed one of these actions to analyse the steps participants take in the decision process. Table 2.11 shows the data we have collected.

The catalog we presented for participants initially presented all laptops, with a short description and a small picture of each one. Additionally, the following actions can be performed in the catalog: (i) *Sort*: laptops can be ordered according to the selected value (price, name, etc.); (ii) *Filter*: different filters (price range, brand, ...) can be added or removed, when the filter links are clicked; (iii) *Show laptop details*: by clicking on the laptop name, a new window opens with the specification of the selected laptop; and (iv) *Compare laptops*: two or three selected laptops can be compared (a table is displayed with laptop features side-by-side).

Table 2.11 shows that the standard deviation of each group is high. It means that, within a group, there are participants that took many more steps to choose laptops than others. Observing the mean value, we see that the participants with lower knowledge levels took more actions to choose their options. When users have little knowledge about the domain, they need to search the catalog to learn about it.

Participants with NO_KNOWLEDGE executed random actions in the catalog, indicating that they had little idea about how to choose the laptop. BEGINNER and INTERMEDIATE participants explored much more to detail laptops in the catalog, showing their exploration of the domain. And ADVANCED and EXPERT

		Steps		
		M	Median	SD
Domain Knowledge	NO_KNOWLEDGE	7.40	1.00	13.35
	BEGINNER	5.56	2.00	9.67
	INTERMEDIATE	3.15	1.00	4.91
	ADVANCED	3.75	2.00	4.46
	EXPERT	4.04	1.00	7.20
Gender	MALE	3.86	1.00	6.41
	FEMALE	3.98	2.00	6.08
Age	16-25	3.57	1.00	6.38
	26-35	4.05	1.00	5.97
	36-45	3.62	3.00	3.71
	>45	4.68	2.00	7.77

Table 2.11: Number of Steps Taken to Choose Laptops.

participants made an extensive use of filters. As they have a more precise idea of what they wanted, they reduced the search space in order to look only at the laptops they were interested in. In case of applications that aid users on the decision process, it is essential to give a personalised assistance that considers their domain knowledge.

In these last three research questions, we do not make any statistical analysis as the standard deviation of some values are very high, for instance, the number of steps taken to choose laptops as seen in Table 2.11. Therefore, we limit ourselves to the qualitative discussion presented above.

2.3 Discussion

In this section, we discuss conclusions we derived from results of our study and present lessons learned that could be used as directions for works that aim at capturing preference specifications from users.

Based on the preference specifications provided by participants, the domain expert was not able to make the choice they made for most of them. However, some participants managed to provide adequate specifications so that the expert made the right choices on their behalf, and in the cases in which this happened, the best choice (first position) was recommended. Therefore, when users provide "good" preference specifications, it is possible to make a decision on their behalf equal to what they would have decided on their own. With the analysis of the specifications, we observed the "good" specifications tend to give the expert orientation about the attributes that matter and preferences over each of them. Nevertheless no information about their interaction (trade-off) has to be given. In general, people use heuristics and principles to resolve trade-offs (Payne et al. 1988), and as the expert tends to use the same principles, in many cases the decision converges into the same choices, made by either the specialist or by the participants. Preferences provided by participants could be given in a short time, regardless of their knowledge about the domain; therefore, it is not unrealistic to expect users to provide preference specifications using a language that is close to natural language.

It is important to highlight one group of participants that made a particular type of specification: no delegation. This group indicated a particular model or explicitly emphasised that they would never delegate such decision to another person (or a system). Therefore, for systems that aim to automate user tasks, it is essential to consider the autonomy degree being achieved and let users control it. Without allowing the user to make the final decision in the decision making process, this group of users will never accept a decision support system.

By analysing changes that participants made after being aware of available options, we observed that most of them did not change their specifications. This indicates that, even though the preference order over available options is constructed when they are actually seen, preferences over individual attributes do not change. When they change, it is because some characteristics were forgotten or they evolved over time and the participant was not aware of it.

As we gave participants the ability to choose more than one option, we could analyse if, after going through the decision making process, they would decide for one and only one option. This was not the case, as participants chose three or four options — and not five, even though it was possible. It shows that people can reduce their search space of options to a very low number, but deciding among them is the most difficult part. In addition, it also shows that there is more than one option that people can accept as an adequate choice.

Finally, it is important to note how participants made the decision. Most of existing work about preferences relies on principles of multi-attribute decision theory (Keeney and Raiffa 1976), and psychology work indicates that people do not follow them (Tversky 1996). If we aim to automate users' tasks, it is important to consider human behaviour, and not only what "rational" decision makers should do, according the definitions of economy. The analysis of the steps that participants with ADVANCED or EXPERT domain knowledge took to make the decision showed that the approach was similar to the one proposed by Tversky (Tversky 1972), namely elimination by aspects, in which people apply cut-off values for attributes according to their relevance.

As discussed above, there is a group of users who are able to express their preferences in such a way that someone can make an appropriate decision on their behalf; and other users need help to specify their preferences. Based on these two groups, we identify different kinds of support for each of them: (a) help for users of the second group to better express their preferences ; and (b) a language that is expressive enough for users of the first group to state their preferences. Next, we present a discussion related to these issues.

2.3.1
Supporting the Preference Expression

Research work on preference elicitation has reported different techniques for it. The kind of support we are looking at is not to elicit preferences from scratch, but to identify issues in preferences specified by users and to help them be more precise. According to the preferences changes of our study, we identified that users do not provide wrong information, but incomplete or out-of-date information in case of values that change over time. In such situations, information about the domain should be provided, such as features left unmentioned, new features and updated values. However, this must take into account the domain knowledge of users so as not to annoy them with things they are aware of. Moreover, some of our participants provided templates of how they specify preferences about laptops, with variables for features that change over time. This can be adopted for providing guidance for users with a starting point for their specification, thus reducing the effort necessary to accomplish this task.

Our study also showed that users typically adopt imprecise adjectives in their preferences statements, even when they are domain experts. A *good* video card has a different meaning for a user who plays games and another who watches movies. Therefore, these adjectives should be identified and scales should be shown to users so they can rate what is "good" or "fast." But the point is to let users express themselves to obtain better specifications later. The same situation happened frequently with the term "cost-benefit." Only one of the participants provided an accurate specification for that: *"5 - the secondary laptops should only be chosen if and only if the price for the same configuration differs around 500 reais, or is 20% or 30% lower than the highest price."* A common issue is also dealing with subjective characteristics, e.g. "modern design," "beautiful." In these cases, samples of groups of items could be shown to users so we could understand what they meant. Naturally, we are not excluding the help of learning algorithms as a complementary approach.

2.3.2
Providing Different Forms of Expressing Preferences

The second point focuses on identifying preference expressions that should be part of a domain-neutral metamodel to represent user preferences, which can be instantiated for different application areas. By analysing the preference specifications of our study, we have concluded that they are significantly different when provided by users with different degrees of domain knowledge. Yet, even

I prefer $\langle target \rangle$	$\langle target \rangle$ is attractive
I (don't) need $\langle target \rangle$	$\langle target \rangle$ is interesting
It is desirable $\langle target \rangle$	I want $\langle target \rangle$
Avoid $\langle target \rangle$	I prioritise $\langle target \rangle$
I (don't) like $\langle target \rangle$	Observe (attribute) $\langle target \rangle$
$\langle target \rangle$ should (not) be A	$\langle target \rangle$ is (not) required
I (don't) want $\langle target \rangle$	I don't care too much about $\langle target \rangle$
$\langle attribute \rangle$ can be $\langle target \rangle$	I make a decision based on $\langle target \rangle$
It is nice to have $\langle target \rangle$	

Table 2.12: Expressive speech acts adopted by participants in assessment statements.

though specifications are indeed different, there is no significant difference in the domain specialist matches among the groups with different knowledge. Therefore, different forms of preference expression must be provided to users, and they are equally important.

We next present common patterns and expressions that we identified in the preference specifications given by the study participants. We group these observations into two main groups, the first associated with preference statements; and the second, preference targets. Preference statements consist of language constructions that indicate preferences over a domain, while targets are the kinds of objects that are referred to by statements. Moreover, we also point out other observed issues, which are related to perlocutionary acts (Back 2006) and trade-offs.

Preference Statements

Five main types of preference statements were identified in the provided preference specifications, which involve *monadic* and *dyadic* preferences. Monadic (classificatory) statements (Hansson 2001) evaluate a single referent, as opposed to dyadic statements, which refer to two referents. The identified types are presented next, and they are associated with one or more targets, which are discussed in next section.

Assessment. In assessment statements, users evaluate a target with a rate or an expressive speech act. We observed that participants, and more generally people, widely use **expressive speech acts**, which include want, need and desire, to express how much they want a certain preference to be satisfied, and we compiled these expressive speech acts in Table 2.12. Although assessment statements refer to a single target, they implicitly establish an order relation among elements, with the information of how much an element is preferred (or equivalent) to another.
Example: *I rate* $< target >$ *with the value* $< rate >$.

Reference Value. Reference value statements enable users to indicate one or more preferred values for an element. These preferred values can also be specified as an **interval**. For instance, there are many occurrences of the specification of an interval for screen size, which is limited by a lower bound, an upper bound, or both.
Example: *I prefer* $< target >$ *as close as possible to* $< reference\ value >$.

Goal. A goal indicates that the user preference is to minimise or maximise a certain element.
Example: *I prefer to maximise* $< target >$.

Order. Order statements establish an order relation between two elements, stating that one element is preferred (strictly or not) to another. A set of instances of the order preferences comprises a partial order.
Example: *I prefer* $< target_1 >$ *to* $< target_2 >$.

Indifference. Indifference statements consist of the indication of a set of elements that are equally (un)important to the user.
Example: I am indifferent to $< target_1 >$ *and* $< target_2 >$.

Many approaches use questions and answers to derive numeric values for user preferences, so that these numbers can be used to calculate a recommendation or make a choice on behalf of the user. However, instead of making users go through a tedious questionnaire, approaches can provide users with the ability of expressing statements such as those presented in Table 2.12. By capturing expressive speech acts used by people, and adopting an interpretation for them, such approaches can detect: (i) which preferences are user requirements (hard constraints); (ii) which attribute values are not the best, but acceptable; and (iii) which attribute values users would appreciate, but are not essential.

With respect to the relative importance among preferences and features, we observed that participants explicitly compared the importance of two features (e.g. “the performance of the laptop is more important to me than its price”) and among feature values (e.g. “I prefer brands A, B and C, in this order”), but several participants also **ordered preference statements**. These participants ranked the provided statements in their specifications, indicating the statements relevance. In addition, the relevance was also expressed by using different expressive speech acts as presented previously. Together, this information is an indication of how much people want to satisfy a particular preference. Moreover, some users explicitly stated which preferences are hard constraints: “*if my preferences are not satisfied, don’t buy it.*” Finally, participants also indicated features that they do not care about.

Preference Targets

Now that we have shown different constructions of preference statements, we focus on their targets. In the literature, often preferences are over a *class* of options (e.g. laptop), which is composed of attributes (or features), and each of which is associated with a domain of values. Given this way of describing a domain, we identified three main preference targets: (i) **class:** I prefer *laptops* to *desktops*; (ii) **attributes:** I don't care about the laptop *colour*; and (iii) **attribute values:** I don't like *laptops* whose *colour* is *pink*. An attribute value is commonly a value of the respective attribute domain, but for some attributes, participants also used subjective values, such as the examples following.

- *Speed*: very slow, slow, normal, fast, very fast.
- *Size*: tiny, small, normal, big, huge.
- *Weight*: very light, light, normal, heavy, very heavy.

Therefore, in many cases, **more than one domain** can be associated with attributes. Moreover, besides grouping attribute values using categories, participants also adopted **adjectives to qualify these attributes**, giving an indication of which values are preferred for an attribute but not being specific, as there is no obvious metrics associated with these adjectives. Examples are: *"comfortable keyboard," "light laptop"* and *"fast laptop."* These adjectives can also be categorically quantified, e.g. *"I prefer a laptop with a screen, whose visibility is **good**."* Furthermore, some of these adjectives are **subjective**, their perception is different for each individual, such as those related to design, quality, reputation and fragility.

Participants can also add new **high-level attributes** (or high-level features) to describe an option, such as a laptop. Participants, mainly those that are not domain experts, tend to express their preferences in terms of high-level attributes about options, being used as a proxy for a set of attributes. For instance, *performance* is related to the processor speed, RAM memory, and so on. The notion of high-level attributes matches the concept of *value*, discussed by Keeney (Keeney 1944). *"Values are what we care about. As such, values should be the driving force for our decision-making. They should be the basis for the time and effort we spend thinking about decisions."* It describes preferences not related to characteristics of the object but the value it brings. Finally, in many situations, participants used a reference option (**prototype**), and express their preferences with respect to it, for instance: *"(second) best model," "most up-to-date," "top configuration"* and *"check the most expensive processor, choose one 10% or 15% cheaper."* This is an interesting way to capture preferences about domains that evolve over time: even though new features of laptops appear constantly, the process of looking for them and stating features relative to a reference point can be used as a pattern for future executions of the task.

Perlocutionary acts

Perlocutionary acts (Austin 1975, Searle 1969) refer to what one achieves by saying something, i.e. the acts that are consequence of saying something. In many cases, when people say something, they mean something else — and this is the situations we report in this section.

First, there are laptops that, generally, are chosen for particular **user profiles**. Therefore, some participants provide their characteristics as an indication of what type of options would be better for them, such as *"I like playing games"* and *"all the places I usually go to are supplied with energy."* Second, participants also indicate which kind of laptop they want by showing for what **purpose** the laptop is going to be used, e.g. *"I want to be able to watch videos on YouTube"* and *"my notebook is a desktop replacement."* Finally, there are cases of **implicit preferences**, as some preferences are common sense, and some participants express them in an implicit way. For instance, there were preference specifications that mentioned: "5. Price," which explicitly indicates that price is relevant. However, it is implicit that the preference is to *minimise* the price.

Trade-offs

Users typically face trade-off situations, such as choosing between a small laptop, with low weight and size, and a big one, with a big screen, but heavy. As the participants did not have available options when they provided their preferences in our study, they did not indicate concrete laptop examples showing the trade-off resolution. The most common expression adopted by participants to indicate how someone should resolve trade-off situations on their behalf is: *"choose a laptop with a good (or the best) cost-benefit relationship,"*, or its variant — *"optimise cost-benefit relationship."* Additionally, few participants provided further details about this relationship, as shown below.

- *"Minimise price"* together with *"I prefer property X, even if that implies a higher price."*
- *"A better brand justifies a higher price at most in 25%."*
- *"5 - the secondary laptops should only be chosen if and only if the price for the same configuration differs around 500 reais, or is 20% or 30% lower than the highest price."*
- *"Not too big a screen because it means a laptop that is too heavy."*

2.4 Final Remarks

This chapter presented a study whose focus is to provide a deeper understanding about user preference specification. We investigated user's ability in expressing their preferences about domains with which they might be familiar with or not, without having experienced a prior decision-making process in such domain with the same options, i.e. users are not aware of the available options. We have targeted the identification of the characteristics of preference specifications provided by users with different degrees of domain knowledge, and how effective they were in order for a domain specialist to use those specifications to make decisions on the users' behalf.

Seven research questions were analysed individually. Our main findings were that users with different knowledge about our study domain, laptops, provide different types of specifications — they are significantly different, according to four types of specifications that we identified as patterns. Users with lower degree of knowledge mainly give high-level preferences and personal information, such as for what purpose the laptop will be used. On the other hand, expert users provide information about fine-grained features. Despite these differences, domain specialists are able to provide recommendations of the same quality for all groups. Therefore, it is essential to provide a rich vocabulary for users to express their preferences, including coarse- and fine-grained preferences. Moreover, we observed users typically provide the right information about their preferences, but they might be incomplete or outdated for preferences whose values evolve over time. This made the domain specialist make a choice on behalf of participants that has at least half of the characteristics of choices made by participants. In addition, mechanisms to help to remember about features to be mentioned and eliminate subjectivity in specifications must be adopted. Finally, we discussed relevant issues to be addressed by preference languages to be used by end-users, such issues include common patterns and expressions of preferences that are typically adopted by people.

It is important to highlight that a limitation of our study is that our findings are based solely on a single domain and the opinion of one specialist, which was due to the lack of other specialists willing to collaborate with our study. However, the knowledge extracted from this study already provides valuable information to propose a domain-neutral metamodel to represent user preferences, which is presented in next chapter. This metamodel takes into account all the identified preference patterns and expressions.