DESIGNING DECISION AIDS FOR DIGITAL SERVICE DESIGN TECHNIQUE SELECTION

Zur Erlangung des akademischen Grades eines
Doktors *der Wirtschaftswissenschaften*

Dr. rer. pol.

von der KIT-Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Xuanhui Liu, M.Sc.

Tag der mündlichen Prüfung: 09. August 2019

Referent: Prof. Dr. Alexander Mädche

Korreferent: Prof. Dr. Gerhard Satzger

Karlsruhe, August 2019

Abstract

Today, people from a multi-disciplinary background are becoming involved in digital service design processes. With the increasing number of digital service design processes in organizations, it is becoming critical to quickly onboard design novices. A huge amount of design techniques is available in digital service design processes. However, too-much-choice makes the selecting process difficult. Thus, selecting appropriate techniques is a challenge, especially for design novices.

This dissertation project focuses on providing decision support for design novices to select design techniques in design processes. Several artifacts in the form of different types of classifications and web-based platforms were developed as decision aids seeking to guide novices to select design techniques in digital service design processes. A design science research paradigm was followed, and three design cycles were conducted in the entire dissertation project. Artifacts were developed and evaluated in each design cycle. In design cycle 1, an expert-based taxonomy and a set of novice-based tags were derived. In design cycle 2, the taxonomy and tags were instantiated as decision aids in a web-based platform and evaluated in a lab experiment. The evaluation demonstrated that the expert-based taxonomy outperformed the novice-based tags. In design cycle 3, an extended version of the web-based platform was developed, including a natural language user interface (UI) in combination with the taxonomy to provide selection support for design techniques. The results of the experimental evaluation of cycle 3 demonstrated that novices' performance using graphical and natural language UI was dependent on decision-making style and duration of use.

The work presented in this thesis contributes to the body of knowledge in the field of digital service design. In design cycle 1, the created taxonomy represents a theory for analysis (type I theory). In design cycle 2 and 3, the proposed design principles represent the core of a theory for design and action (type V theory) to guide the design of decision aids for supporting design novices' selection of design techniques. Besides the theoretical contribution, the developed platform for the selection of design techniques contributes practically to help design novices select design techniques under different design situations.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
DF	Design Feature
DP	Design Principle
DSR	Design Science Research
DSS	Decision Support Systems
EFA	Exploratory Factor Analysis
IPO	Inputs-Process-Outputs
IS	Information Systems
IT	Information Technology
MR	Meta-requirement
RQ	Research Question
SD	Standard Deviation
SUS	System Usability Scale
UI	User Interface
UX	User Experience

1 Introduction[1]

With the wide adoption of the Internet and mobile technologies, more and more traditional face-to-face services are gradually being replaced by digitalized services (Schneider 2017). In order to successfully design digital services, plenty of design techniques have been developed to help in conducting design activities in design processes (e.g., Curedale 2013; Martin and Hanington 2012; Stickdorn and Schneider 2011). However, as more choices lead to challenges in the decision-making process, the selection of suitable design techniques is difficult (Bollen et al. 2010; Greifeneder et al. 2010), especially for design novices who have limited knowledge (Bonnardel et al. 2003). Therefore, selecting appropriate design techniques from the broad spectrum of possibilities becomes a challenge. Hence, decision aids are needed for supporting design novices in the selecting process of design techniques.

1.1 Motivation

The digital world witnesses a shift in the relations between companies and customers from one-time transactions to ongoing relationships in which customers continuously use digital services (Vargo and Lusch 2004). With the transition from selling services to selling experiences, the design of digital services is getting even more challenging (Deloitte Digital 2019; Mager 2009; Pine and Gilmore 1998). In order to provide services with high usability and user experience (UX), more and more organizations recognize the importance of design-oriented approaches (Deloitte Access Economics 2015; Gartner 2019). Take Gartner's statement as an example, in recent years, in digital service-oriented companies such as Netflix, Google, Facebook, etc., the ratio of designers to developers was between 1:5 and 1:4 (Revang 2017). However, the self-reported average ratio of designers to developers from the companies of Gartner clients was 1:17, which is considered as an improper amount of designers (Revang 2017). Thus, design novices are also involved in design processes. The process of digital service design combines perspectives from business, technology, and design, and therefore, the entire service environment needs to be considered in the design process (Buchanan and McMenemy 2012; Patrício et al. 2011), people with multi-disciplinary backgrounds are engaged in design processes (Evenson 2005). For example, product owners in agile software development teams are expected to systematically leverage techniques from user research, while software developers are encouraged to learn and

[1] This Chapter is based on the following studies which are published or in work: Liu et al. (2016); Liu, Leung, et al. (2018); Liu (2018); Liu, Werder, et al. (2018); Liu, Werder, et al. (2019a); Liu, Werder, et al. (2019b); Liu, He, et al. (2019); Liu, Rietz, et al. (2019).

apply techniques to test usability in order to provide user-friendly software (Bruun 2010). There is an increased need for assisting people in selection and use of design techniques while developing digital services.

As design processes can benefit from systematic guidance, a broad spectrum of design techniques has been developed. The design techniques are intended for various situations, such as generating design ideas, creating high-fidelity prototypes, or collecting user feedback. In the context of digital service design techniques, one can distinguish three research streams: i) the development of design techniques (e.g., Schleicher et al. 2010; Tähti and Niemelä 2006), ii) the application of design techniques (e.g., Judge et al., 2008; Vanattenhoven and De Roeck, 2010), and iii) the classification of design techniques (e.g., Curedale 2013; Martin and Hanington 2012). Research on development introduces newly developed design techniques, such as the creation of touchpoint matrix (Brugnoli 2011), experience clip (Isomursu et al. 2004), and experience prototyping (Buchenau and Suri 2000). Research on the application of design techniques explains their use and purpose, such as using Kano analysis to understand user needs (Wang and Ji 2010) or using affinity diagramming in a group discussion to identify user needs (Judge et al. 2008). Research on classifying techniques investigates the guidance for enhancing team coherence (Alves and Nunes 2013) and using design techniques for different situations (Maguire 2001).

However, with the expansion of the number of design techniques, more alternatives exist in the selection process. More alternatives can cause more difficulties when selecting the desired design techniques for specific design situations (Bollen et al. 2010; Greifeneder et al. 2010), especially for design novices who have limited experience and knowledge (Bonnardel et al. 2003; Sanders et al. 2010). Yet, no existing research focuses on systematically providing decision support for design novices to select design techniques.

Classifications describe the similarities and differences of the categorized items, and can be seen as a basis for locating the required items (Bailey 1994). Classifications have been widely applied, e.g., in biology and medicine. In addition, classifications typically reflect the structure and hierarchy of a website, which helps users to find the information they need (Boulton et al. 2016); for example, helping job seekers to find appropriate occupations (ESCO 2017). In e-commerce, customers can select products based on their preferences and the product attributes (Xiao and Benbasat 2007), which are classified into pre-defined categories and can be used to reduce the number of alternatives (Wang and Benbasat 2009). Thus, building a classification can be seen as a starting point for providing a decision aid. Although there is research on classifying design techniques, existing classifications of design techniques lack rigor and completeness. Examples identified in the existing literature deliver ambiguous results and lack empirical evidence for their conceptualization (e.g., Alves and Nunes, 2013; Roschuni et al., 2015b). In addition, the existing classifications of design techniques are built based on experts' knowledge (e.g., Roschuni, Kramer, and Agogino 2015; Roschuni, Kramer, Qian, et al. 2015). However, experts and novices have different categorization behaviors (Chi et al. 1981; Mitchell 1989). Experts tend to build classifications with hierarchical categories from a top-down manner, while novices incline to classify entities into flat categories without any parent-child relations by following a bottom-up approach (Chi et al. 1981; Mitchell 1989). Novice-based classifications have a wide application, for example, user-knowledge based tag clouds are usually used for searching information in website navigation (Klašnja-Milićević et al. 2017). Hence, novices' understandings of design techniques should be considered when creating classifications. Furthermore, the effect of classifications on the selection of design techniques has not yet been evaluated in the literature. Although classifications are studies that can help with the selection for certain purposes (Glass and Vessey 1995), how much can classifications help novices to select design techniques is still blank at present.

Moreover, the presentation of classifications can be different, such as a structured navigation menu (i.e., a taxonomy-based graphical UI) (Bussolon 2009) or a conversational agent (i.e., a natural language-based UI) (Chen et al. 2017). The different visualization of a classification can also influence the performance of classification. A hierarchical taxonomy uses abstract terms to describe the categories (Nickerson et al. 2013), while a natural language dialog can explain the abstract terms in an easier way (McTear 2002). Existing research has investigated natural language generation to explain the categories of a classification (Wulf and Bertsch 2017). However, no research compares the performance of the taxonomy-based graphical UI and the natural language-based UI.

With my doctoral research, I seek to close this practically and scientifically relevant research gap by suggesting and systematically evaluating different forms of decision aids in order to support design novices to select design techniques.

1.2 Research Questions

Based on the theoretical and practical motivations, this doctoral research seeks to answer the following research question:

Main RQ: *How to support design novices in digital service delivering organizations to select digital service design techniques?*

In order to answer the main research question, the literature from the three research domains: classification, design techniques, and decision-making were reviewed and analyzed. This generated three break-down research questions to answer the main research question step-by-step. An overview of the three research domains and research gaps is presented in figure 1-1.

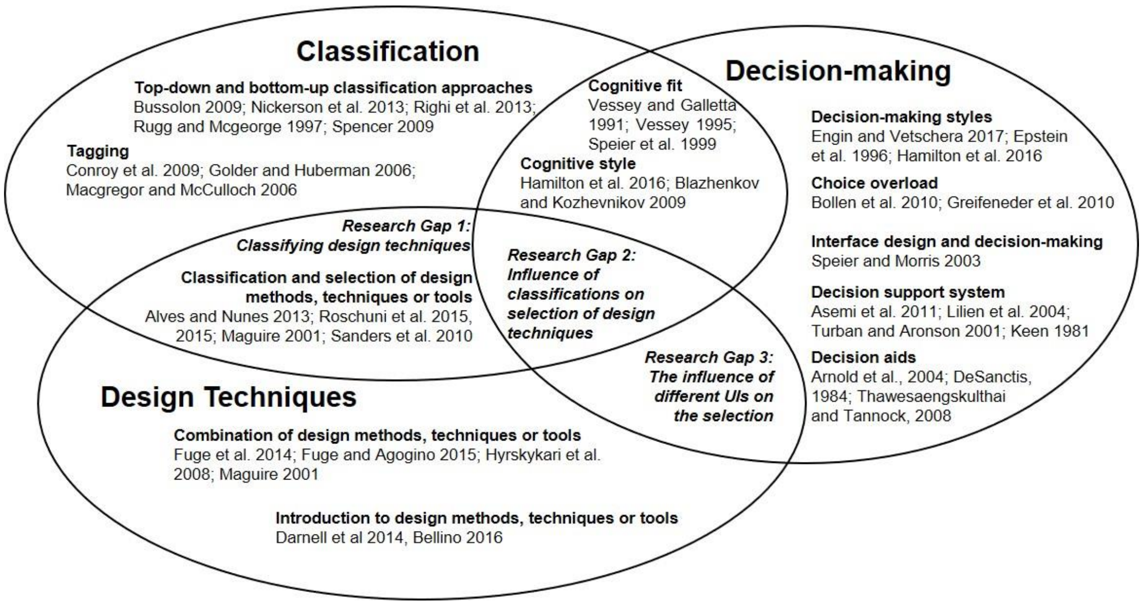


Figure 1-1. Overview of research domains and identified research gaps

As the purpose of this study is to help design novices in the selection process, not only design experts, but also design novices' understandings of classifying design techniques need to be considered. There are differences between experts and novices when classifying items because of the differences in their knowledge of the specific domain. Experts tend to build classifications from a top-down perspective, while novices create classifications using a bottom-up approach. Thus, classifications from both perspectives need to be provided. The first research gap is about classifying design techniques using top-down and bottom-up approaches from the perspective of design experts and novices. Hence, the first break-down research question is:

RQ 1: *What categories build classifications of design techniques from both experts' top-down and novices' bottom-up perspectives?*

By answering ***RQ 1***, an expert-based taxonomy (top-down) and a set of novice-based tags (bottom-up) are developed. Although classifications can help with depicting similarities as well as differences and reducing the complexity, they cannot be directly used to support the selection of design techniques. Thus, in order to use the classifications in the selection of design techniques, a web-based platform that instantiate the two classifications needs to be provided. These two types of classifications (i.e., hierarchical taxonomy and tags) are widely used in websites as a decision aid for finding information (Sinclair and Cardew-Hall 2008). With the web-based platform, the selection process of design techniques can be supported by narrowing the selection scope of design techniques. In other words, instead of selecting design techniques from a large amount, the selection scope can be narrowed by selecting the relevant categories before looking into the detailed description of each design technique. However, no research analyzes or compares the performance of the two types of classification. It is interesting to see whether novices can benefit more from an expert-based classification (i.e., top-down taxonomy) or a novice-based classification (i.e., bottom-up tags) in the selection process. In addition, individual differences of decision-making styles influence the decision-making process (Hamilton et al. 2016; Lee et al. 2007); rational decision-makers are analytic and tend to structure information and compare them before making decisions, while intuitive decision-makers are holistic and tend to follow their hunch to make decisions. Thus, using different types of classifications for a decision-making task also needs to consider decision-making styles. Based on the cognitive fit theory, the match between problem representation and the problem-solving task can influence task performance (Vessey 1991). As design novices have limited knowledge of design techniques, the match between the understanding of design novices and the provided decision aids also needs to be considered. Thus, an evaluation of the influence of different classifications in the selection process of design techniques needs to be conducted. The second research gap focuses on the instantiation of classifications and the analysis of the effect of different instantiations on the selection performance. Hence, the second break-down research question refers to:

RQ 2: *How to instantiate the classifications and how can classifications influence the design novices' selection of design techniques with the consideration of decision-making style?*

By answering ***RQ 2***, the results present that the expert-based taxonomy generally outperforms the novice-based tags when selecting appropriate design techniques for specific design scenarios. As the terms of the categories in the taxonomy are abstract and mostly not self-explained, a follow-up study of the taxonomy is to present the abstract content of the expert-based taxonomy in an easier way using a natural language dialog to guide novices to select design techniques (Maes 1994). Examples of using natural language dialogs for searching information can be found in many websites, besides a hierarchical navigation menu with abstract terms, a chat frame with natural language dialog is provided to help people find the desired information. Although the categories of the information are the same, users may perform differently when searching for information because of the different UIs (Speier and Morris 2003). Yet, there is no research on the effect of different UIs of taxonomy on the selection process. In addition, as rational decision-making styles influence the processing process of structured information, rational decision-making style is considered when comparing different UIs of the hierarchically structured taxonomy. Thus, the third break-down research question is:

RQ 3: *What kind of UI presentation of a taxonomy of design techniques can benefit the novices' selection of design techniques with the consideration of rational decision-making style?*

Figure 1-2 presents an overview of the three break-down research questions and the main research question. With the three break-down research questions, the main research question was answered step-by-step. There was a progressive relationship between the three research questions.

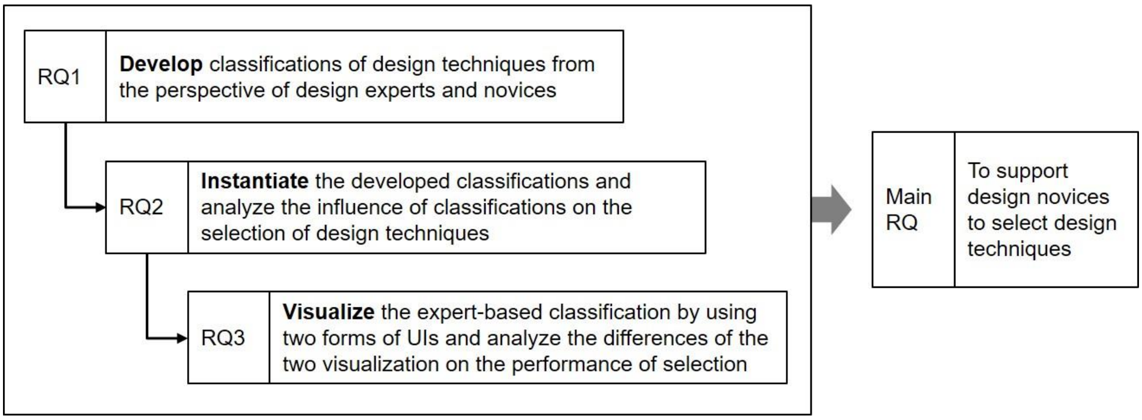


Figure 1-2. Overview of research questions

1.3 Dissertation Structure

This dissertation is structured based on the design science research (DSR) paradigm (Kuechler and Vaishnavi 2012) and organized in three cycles. Figure 1-3 presents an overview of the dissertation structure.

In Chapter 2, the theoretical foundations and related work are introduced. In Chapter 3, the overall DSR methodology with three design cycles is described. The first design cycle is presented in Chapter 4 and explains the derived classifications of design techniques from a top-down and bottom-up perspective. RQ1 was addressed in two steps. Part 1 (Chapter 4.1) describes the development of a top-down taxonomy from the perspective of design experts. Part 2 (Chapter 4.2) describes the development of bottom-up tags from the perspective of design novices. Chapter 5 presents the second design cycle which instantiates and evaluates the developed classifications in cycle 1. RQ2 was answered by dividing cycle 2 into two parts. Part 1 (Chapter 5.1) introduces the development of a web-based platform that instantiates the taxonomy and tags created in cycle 1. Subsequently, in part 2 (Chapter 5.2), an evaluation of the effects of the taxonomy and tags on novices' selection performance when selecting design techniques is described based on a lab experiment. Chapter 6 is again structured into two parts. Part 1 (Chapter 6.1) introduces a refined version of the web-based platform with an attempt to provide an alternative presentation of the taxonomy using a natural language dialog. In order to answer RQ3, the comparison of the effect of using a taxonomy-based graphical UI and a natural language-based UI in a lab experiment is explained in part 2 (Chapter 6.2). Chapter 7 discusses the contribution, limitation, and outlook of this doctoral project. Finally, Chapter 8 concludes the dissertation.

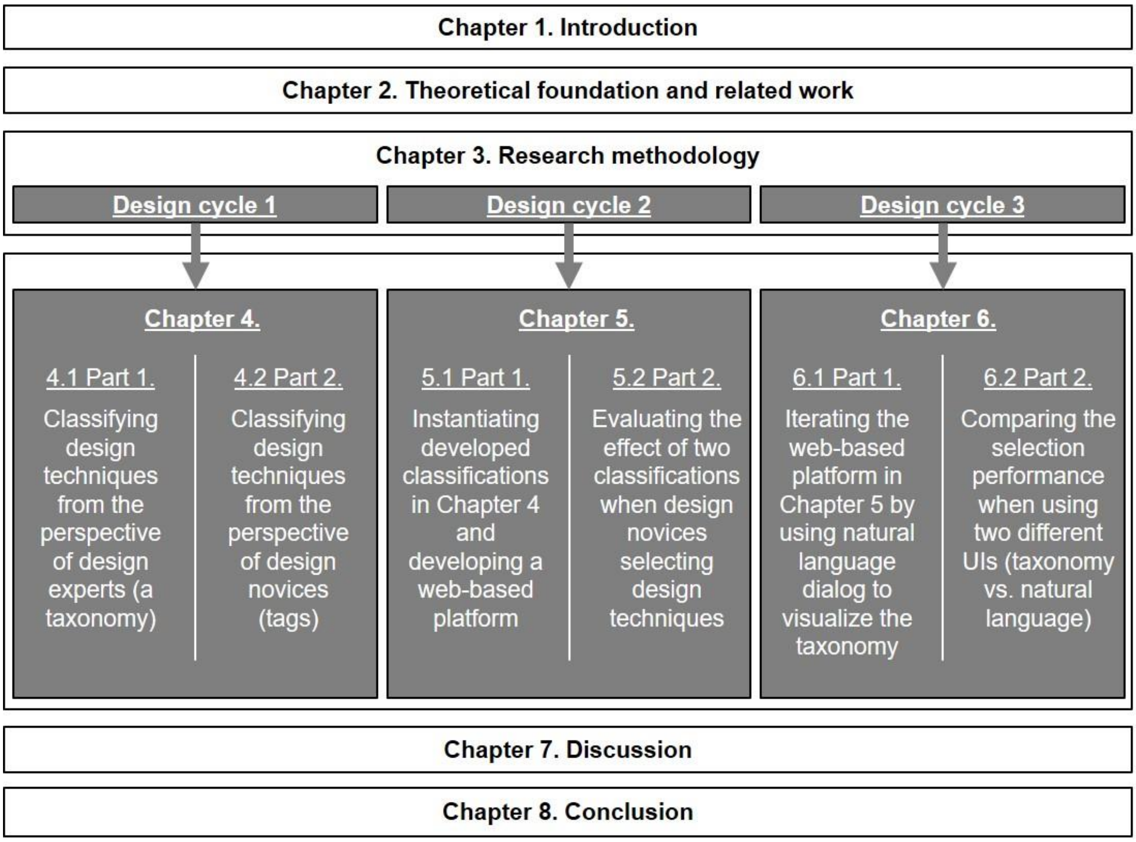


Figure 1-3. Overview of the structure of this dissertation

2 Conceptual Foundations and Related Work[1]

In this Chapter, the conceptual foundations and related studies are introduced as a background for this dissertation. Digital service design, design techniques, and selection support are explained and conceptualized. As the focus of this dissertation is to support design novices when selecting design techniques, a comparison between design experts and novices is described. In addition, the theoretical concepts of cognitive fit and decision-making style are described. Subsequently, the related work is introduced from both a scientific and practical perspective.

2.1 Digital Service Design

Digital service is defined as *"an activity or benefit that one party can give to another, that is, provided through a digital transaction"* (Williams et al., 2008, p. 507). Designing digital services differs from designing non-digital services in several ways. A personal relationship is a basis for non-digital service, whereas in the context of digital services face-to-face contacts are secondary (Williams et al. 2008). The interaction between a human and a digital service can be realized by websites, mobile applications, etc., which transforms traditional human-based services into human-free services, such as ticket sales, online check-in, or online banking (Fan et al. 2013; Sakata et al. 2014; Spagnoletti et al. 2015). The design quality of a digital service determines user perceptions, such as satisfaction, usability, and UX (Barrera-Barrera et al. 2015; Cardoso et al. 2015; Collier and Bienstock 2006; Ozen Seneler et al. 2010). Moreover, some studies suggest that supra-functional needs (e.g., emotional, aspirational) are more important than functional needs (Williams et al. 2008). People who participate in digital service design processes are expected to create a seductive UX (Gruber et al. 2015).

Digital service design processes include many activities, e.g., comprehensive user behavior analysis, creation, and evaluation of prototypes. As digital service design processes emphasize human-free interactions, user habits need to be investigated rigorously, and design results need to be tested several times before a release and will be further optimized after release (Gruber et al. 2015). In iterative design processes, evaluation activities are not limited to the end of the design process (summative evaluation), but are also performed through the entire design process (formative evaluation). Evaluation and design activities are conducted in parallel and studied

[1] This Chapter is based on the following studies which are published or in work: Liu et al. (2016); Liu, Leung, et al. (2018); Liu (2018); Liu, Werder, et al. (2018); Liu, Werder, et al. (2019a); Liu, Werder, et al. (2019b); Liu, He, et al. (2019); Liu, Rietz, et al. (2019).

jointly during the iterative process (Brhel et al. 2015; Spaulding and Weber 2009). Designers conduct activities to understand users, create ideas, and analyze outcomes to choose the most appropriate idea and find the best solution (Werder et al. 2016). While some activities emphasize the creation of ideas, such as drawing storyboards (Henrikson et al. 2016) or sketches (Tohidi et al. 2006), other activities help to compare different design ideas, such as A/B testing (Kohavi et al. 2007). Hence, when conducting activities in digital service design processes, design and evaluation are complementary and mutually reinforcing. In the following, the term "design" is used when referring to both of the creation of prototypes or sketches as well as the analysis of user habits or the evaluation of ideas.

2.2 Design Methods, Techniques, and Tools

In digital service design processes, various activities need to be applied to achieve design goals. Besides identifying the context, people need to understand the uniqueness of usability work, which needs the combination of insights, experience, judgment, and wisdom in the approach (Woolrych et al. 2011). The successful creation and application of design activities need experience, time, focus, and attempts, which is built on the understanding of a comprehensive overview of design methods, techniques, and tools. While the three terms - methods, techniques, and tools - are rather ambiguously used in practice, they have distinct definitions from a conceptual point of view. Figure 2-1 presents the relationship between the three terms.

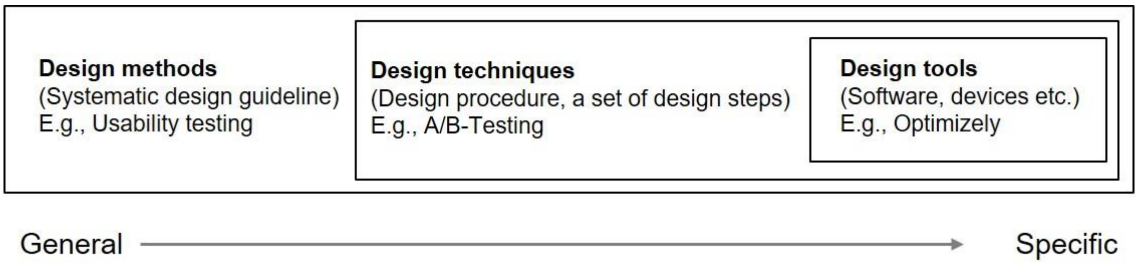


Figure 2-1. Design methods, techniques, and tools

A design tool is an instrument to support the design process, for example, a software, an analysis form, a device (Hackathorn and Karimi 1988; Palvia and Nosek 1993). The development of new design tools evolves fast, and the application of a tool is rather specific. A tool can either support a couple of different design techniques or a whole design process (Brinkkemper 1996). A design technique can be defined as a procedure or a set of steps to achieve desired outcomes (Brinkkemper 1996; Hackathorn and Karimi 1988; Kettinger et al. 1997). In comparison with a tool, a technique is more generic. After selecting an appropriate design technique, people will have a limited number of design tools to choose from. For example, when choosing A/B testing to compare two designs of a website (Kohavi et al. 2007), a designer can select a specific tool, such as Optimizely to collect and analyze data (Dumas and Salzman 2006). A design method provides an approach for carrying out design activities, which needs systematic and specific ways of thinking and consisting of rules (Brinkkemper 1996). For example, a usability test is a method. When carrying out a usability test, many techniques and tools are needed, such as retrospective think-aloud or a SUS questionnaire (Al-Wabil et al. 2010; Brooke 1996). Tools and techniques are strategically combined in a method, which gives a method a broader scope than a tool or a technique (Sanders et al. 2010). Based on the descriptions of the definition of design methods, techniques, and tools, this doctoral project focuses on classifying design techniques in this study. As mentioned above, the difference between methods and techniques in the context of digital service design is not always consistent, e.g., heuristic evaluation or cognitive walkthrough are sometimes called methods and sometimes techniques (Bruun 2010; Hertzum and Jacobsen 1999; Jeffries et al. 1991; Rieman et al. 1995). As a design technique contains a procedure for conducting the activities, it is easy for design novices to use in the design process by following the steps introduced in a design technique (Bonnardel et al., 2003).

2.3 Selection Support

Decision support systems (DSS) provide decisional advices (Turban and Aronson 2001) and implemented in various domains, such as marketing, management, etc. (Asemi et al. 2011; Lilien et al. 2004), which can benefit decision-makers in many ways: an increase in the number of alternatives examined, a better understanding of the business, fast response to unexpected situations, improved communication, cost savings, better decisions, better use of data resources, etc. (Keen 1981). The focus of this dissertation is to design decision aids to help the selection of design techniques, which does not include all the features of DSS but emphasizes on the selection support. In this Chapter, decision aids, classifications, and natural language UI of classification are introduced.

2.3.1 Decision Aid

Decision aids are defined as interventions to help people make choices among options by providing information about the options and outcomes. Decision aids are intended to improve decision quality (Elwyn et al. 2010) by supporting the selection process leveraging decisional guidance or explanations (Messier 1995). There are many different forms of decision aids, such as graphics, documents, tables, checklist, classifications, etc. (e.g., Arnold et al. 2004; DeSanctis 1984; Thawesaengskulthai and Tannock 2008). Decision aids have shown to positively influence the decision-making process in many studies. For example, rules are used for selecting requirement engineering techniques (Jiang et al. 2008); multi-criteria approaches are used for selecting management methodologies (Thawesaengskulthai and Tannock 2008); interactive decision aids are used to support customers' decision-making in e-commerce (Wang and Benbasat 2009). When using a decision aid, system cognitive efforts are needed to understand the decision aid in order to use it appropriately (Johnson 2010). System cognitive effort means the effort people spend on processing information and understanding the system when conducting a task (Pereira 2000). Moreover, if the decision aid is beyond users' capability to understand the content, users might be misled during the decision-making process (Arnold et al. 2004). For example, a study on intelligent decision aids shows that novice decision-makers may tend to make poor decisions when the decision aids need more expertise than novices already have (Arnold et al. 2004).

2.3.2 Classification, Taxonomy, and Tags

The term classification is defined as *"the general process of grouping entities by similarity"* (Bailey 1994). As a form of decision aid, a classification can be created with hierarchical structured dimensions and categories (i.e., taxonomy) (Nickerson et al. 2013) or with flat categories (i.e., tags) (Mathes 2004). In the field of information architecture, it is known that both expert-based taxonomy and novice-based tags can help with navigating, filtering, and selecting the desired information (Golder and Huberman 2006). Figure 2-2 depicts the differences between the structure of taxonomy and tags.

Taxonomy is a form of classification (Wand et al. 1995), which is defined as *"the theoretical study of the classification, including its bases, principles, procedures, and rules"* (Simpson 1961). Similar to a multidimensional classification, a taxonomy includes dimensions and characteristics (Nickerson et al. 2013). Take classifying animals as an example to explain taxonomy, animals can be classified from the dimension of the food they eat (e.g., herbivore, carnivore) or from the dimension of morphology (e.g., birds, mammals). In this example, the classified animals are entities, and dimensions can be understood as criteria to classify the entities, while the groups (e.g., herbivore, carnivore) of animals under dimensions are categories (Rugg and McGeorge 1997). Relying on domain experts' knowledge, a taxonomy with high-level dimensions and low-level characteristics can be developed (Goh et al. 2009). A taxonomy identifies similarities and differences between entities and reduces the complexity, which is widely applied in various aspects of life (Bailey 1994). For example, taxonomies can be used to help with the navigation of a website (Broder 2002). The hierarchically structured dimensions and characteristics can be used as filtering conditions when searching for information.

Besides a structured taxonomy, so-called folksonomies (i.e., tags) based on an unstructured knowledge organization paradigm (Klašnja-Milićević et al. 2017) can be used. Tags are created by *"grassroots community,"* including no parent-child relations (Mathes 2004), which represent users' preferences of the terms for the categorization (Peters and Stock 2008). Taking the same example of classifying animals, unlike a taxonomy with categories structured under certain dimensions, tags are not hierarchically structured. The categories of animals such as mammals, birds, herbivores, etc. are at the same level without any high-level dimensions. Although the tags are flat, the used term in the tags may have a common understanding within the user community (Goh et al. 2009; Marlow et al. 2006). Users also benefit from using tags. For example, tags enable users to collect resources, organizing information, search information and retrieving them (Golder and Huberman 2006; Marlow et al. 2006; Voss 2007).

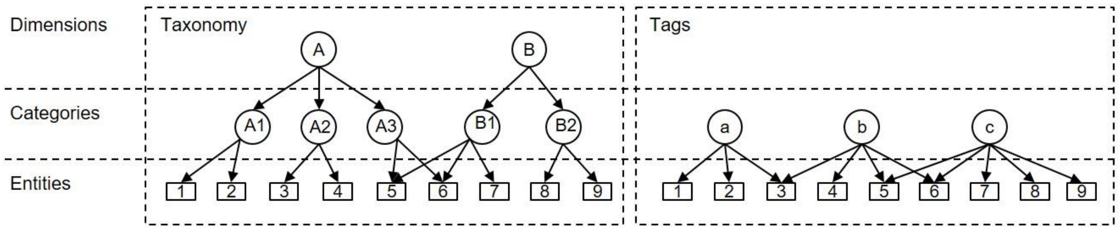


Figure 2-2. Differences in the structure of taxonomy and tags

2.2.3 Taxonomy-embedded Natural Language Dialog

Although the dimensions and characteristics in the taxonomy can be seen as conditions when filtering information, the terms in the taxonomy are abstract (Nickerson et al. 2013), which might be difficult for novices to understand. A natural language dialog system can be built on the taxonomy, which explains the dimensions and characteristics as pre-defined questions (Chen et al. 2017). Natural language dialog systems can be classified into three categories (Arora et al. 2013; McTear 2002): finite state-based system, frame-based system, and agent-based system. In the type of finite state-based system, the dialog is based on pre-defined sequences of steps or stages. The dialog is easy to construct. The system verifies users' answers in each state of the dialog (McTear 2002). But user inputs are restricted to single words and not allowed to ask questions or take the initiative (Arora et al. 2013). An example of such a dialog was for example used in early travel planning systems (Bobrow et al. 1977). The system extracts keywords from users' input to fill out a form. In the type of frame-based system, the dialog is not fully pre-defined but based on the user input (Arora et al. 2013; McTear 2002). Compared with the finite state-based system, the frame-based system allows a more natural dialog. However, the system still cannot process complex inputs. For example, Allison (2012) provides a conversational agent to answer questions about library resources. The conversational agent uses a pre-defined structure and organized data. The system answers users' questions by matching user inputs with the categorized information in the database. In the type of agent-based system, the communication is interactive between the conversational agent and users, and the dialog can evolve dynamically (Arora et al. 2013; McTear 2002). However, as user inputs are highly flexible, the result of the dialog can be incorrect and misleading when the system matches the user input with the wrong data (Hepenstal et al. 2019).

2.4 Design Experts and Novices

There are many differences between experts and novices. They have been studied in various fields, such as biology, physics, and information systems (IS) (Chi et al. 1981; Galletta and Dunn 2014; Nah and Benbasat 2004; Schenk et al. 1998; Shafto and Coley 2003). A key difference is the availability of a domain-specific knowledge base, which is a basis for problem-solving. Compared with experts, novices have limited domain-specific knowledge (Kolodner 1983). As a result, novices are not as skilled as experts when relating problems to their knowledge and previous experiences (McKeithen et al. 1981; Schenk et al. 1998). Novices need more efforts than experts to access information and find suitable solutions (Schenk et al. 1998). The knowledge base of novices limits the abilities for problem-solving (Schoenfeld and Herrmann 1982). When solving problems, experts consider multiple methods, while novices may use the first method that comes to their minds based on the superficial features of the problems (Chi et al. 1981; Jeffiies et al. 1981). An example can be found in the study of the differences between expert and novice programmers. Compared with expert programmers in performing debugging tasks, novices need more time and tend to use breadth-first approaches without formulating an overall program structure model (Vessey 1985). The different availability of a domain-specific knowledge base leads to differences in the organization of the knowledge base (Chi et al. 1981). When conducting categorization, experts tend to build hierarchical structures, while novices incline to create a classification with a set of flat (i.e., non-hierarchical) categories (Chi et al. 1981). Additionally, the labeling of the categories is different. Experts label categories based on principles in a conceptual form, while novices name the categories based on superficial features. For example, an experiment conducted by Shafto and Coley (2003) for classifying marine creatures presents that novices mostly use the appearance of marine creatures as a basis for categorizing marine creatures, whereas experts use environmental or behavioral factors to justify sorts. The differences between experts and novices may also exist in classifying design techniques.

Bonnardel et al. (2003) use the term “lay-designers” to describe novices in the field of web page design. *“Lay-designers are people with little or no formal training in either web design specifically or its attendant skills (e.g., graphic design, user interface design, etc.)”* (Bonnardel et al., p. 26). Design novices in usability engineering are defined as *“persons with less than one year of job experience and no formal training in usability engineering methods”* (Bruun, 2010, p. 83). Thus, design novices are defined in this dissertation as persons with no formal training in digital service design activities. Because of limited knowledge, design novices may have difficulties when selecting design techniques, especially when there are a lot of design techniques out there.

Based on the introduction of selection support, as taxonomy and tags can both help with reducing complexity and distinguishing the similarities and differences, design novices can benefit from using taxonomy or tags when searching for desired information (Golder and Huberman 2006; Liu et al. 2016). Except for these two explorative ways, getting suggestions from experts is more direct. However, as many organizations lack specialized design experts (Bruun 2010), a natural language UI embedded with categorized techniques can be provided (Chen et al. 2017).

2.5 Decision-making Style and Cognitive Fit

The analysis of break-down research questions identifies the use of classifications for searching desired information needs the consideration of different decision-making styles. Thus, two main decision-making styles that can influence the selection process and cognitive fit are introduced.

2.5.1 Decision-making Style

Decision-making styles are based on cognitive styles, which can be roughly distinguished as analytical and intuitive (Allinson and Hayes 1996; Epstein et al. 1996). Cognitive styles are defined as *“an individual’s preferred way of gathering and processing data”* (Allinson and Hayes 1996). The differences between rational information processing and intuitive information processing are compared in 12 dimensions by Epstein et al. (1996) (Table 2-1). The comparison of the 12 dimensions includes effort used, logical based, emotional control, etc. in processing information when people are taking actions and making decisions. Different cognitive styles lead to different decision-making styles when making decisions (Kozhevnikov 2007). As this dissertation focuses on decision-making, the dimensions (e.g., information reasoning logic, encoding, and connection) that are relevant to decision-making are analyzed in more detail.

Table 2-1. Comparison between rational and intuitive style, from Epstein et al. (1996)

Rational	**Intuitive**
• Analytic • Intentional, effortful	• Holistic • Automatic, effortless
• Logical: reason oriented	• Affective: pleasure-pain oriented
• Logical connected	• Associated connected
• Encodes reality in abstract symbol, words, and numbers	• Encodes reality in images, metaphors. And narratives
• Slower processing: oriented toward delayed action	• More rapid processing: oriented toward immediate action
• Changes more rapidly and easily: changes with strength of argument and new evidence	• Slower and more resistant to change: Change with repetitive or intense experience
• More highly differentiated	• More crudely differentiated: Broad generalization gradient; stereotypical thinking
• More highly integrated: Context-general principles	• More crudely integrated: Dissociative, emotional complexes; context-specific processing
• Experienced actively and consciously: We are in control of our thoughts	• Experienced passively and preconsciously: we are seized by our emotions
• Requires justification via logic and evidence	• Self-evidently valid: "Experiencing is believing"

Rational decision-makers tend to gather and compare all the information before making the final decision (Hamilton et al. 2016). Intuitive decision-makers incline to trust feelings and make decisions fast (Hamilton et al. 2016). In a decision-making task, people with a rational decision-making style are analytic and based on evidential factors such as systematically gathering information as well as considering the alternatives, whereas people with intuitive decision-making style are holistic and based on feelings (Epstein et al. 1996; Hamilton et al. 2016). The rational decision-making style can influence the processing of structured information. For example, people with rational cognitive style tend to understand and search structured information when using self-service technologies (Simon and Usunier 2007). As the information categorized in a taxonomy is hierarchically structured, which corresponds to the rational decision-making style, the rational decision-making style might influence the use of a taxonomy. In addition, the structured information can be visualized in different ways (i.e., taxonomy or natural language UI). Thus, processing the hierarchically structured information included in different UIs might also be influenced by rational decision-making style. In contrast to the rational decision-making style, intuitive decision-making style leads to an unstructured and randomized information processing process in a decision-making task (Epstein et al. 1996). Similarly, the creation of tags is also unstructured (Mathes 2004). Thus, intuitive decision-making style might influence the use of tags.

2.5.2 Cognitive Fit Theory

The cognitive fit theory proposes that the problem-solving performance improves when the problem representation matches the task; and matching the problem-solving skill to the problem representation and task affects performance (Vessey and Galletta 1991) (Figure 2-3). The theory was used to analyze people's performance when using a graphical or tabular format of information for solving problems (Vessey 1991; Vessey and Galletta 1991). The theory is widely cited (1264 times till April 7th, 2019 from Google Scholar) and applied in studies relating to decision-making task (e.g., Huysmans et al. 2011; Speier et al. 2003). The problem representation refers to the system that provided for the decision-making task. The problem-solving task is a decision-making task. The problem's solution is the individual task performance on the given task by using the provided system. Mental representation is generated when using the provided system to solve the decision-making task.

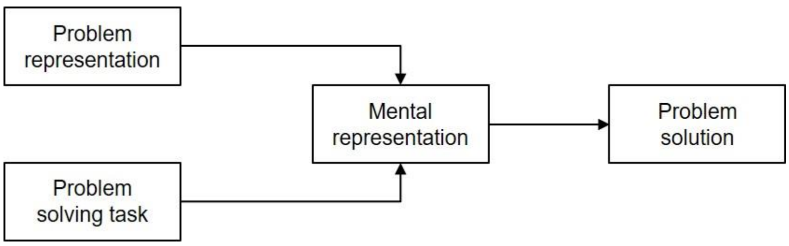


Figure 2-3. Cognitive fit mode, from Vessey (1995)

The cognitive fit theory was extended by splitting problem representation into internal and external problem representation. The internal problem representation is measured by self-evaluation of the domain knowledge and experiences (Mangalaraj et al. 2014; Shaft and Vessey 2006). The external problem still means the system provided for solving the task (Shaft and Vessey 2006; Vessey 1991). Both internal and external representations contribute to the mental representation (Shaft and Vessey 2006). A mismatch between external representation and mental representation can negatively influence the effectiveness and efficiency. The extended cognitive fit was, for example, applied in the data exploration analysis of visual representation (Baker et al. 2009), design patterns and collaboration in software design (Mangalaraj et al. 2014), etc.

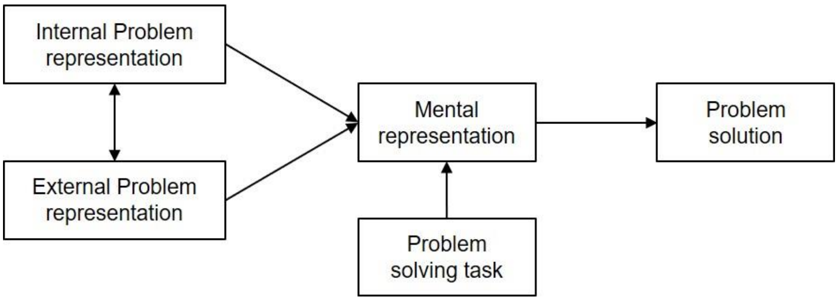


Figure 2-4. Extended cognitive fit mode, from Shaft and Vessey (2006)

2.6 Related Work

In order to reduce the complexity of the selection caused by the increasing number of design techniques, several studies proposed classifications of design techniques. Additionally, some of the developed classifications are instantiated as web-based systems from a more practical perspective.

2.6.1 Related Work in Design Technique Classifications

Various fields, such as design thinking (Roschuni, Kramer, Qian, et al. 2015), service innovation (Kumar and LaConte 2012), human-centered design (IDEO 2015; Maguire 2001), participatory design (Sanders et al. 2010), and service design (Alves and Nunes 2013; Curedale 2013) deal with the challenge of designing digital artifacts using different design techniques. Thus, the idea of providing a classification for design techniques is not entirely new. Table 2-2 summarizes important studies and describes their field of origin, the purpose, the main dimensions, and research approaches of previous studies in this context.

Table 2-2. Related work on design technique classifications

Study	Field	Purpose	Dimensions	Research Approach
(Roschuni, Kramer, Qian, et al. 2015)	Design thinking	To provide a standardized way for the communications of design methods	Design process (research, analysis, ideate, build, communication)	Workshops and theoretical studies
(Kumar and LaConte 2012)	Service innovation	To engage in successful design innovation	Benefit, input, output, innovation process	-
(Martin and Hanington 2012)	Design method	To support the design of meaningful products	Type of content, form of content, phases, origin, roles of researchers, primary purpose	-

(IDEO 2015)	Human-centered design	To provide a field guide on implementing human-centered design	Design process, difficulty, time, participant	Case study
(Maguire 2001)	Human-centered design	To guide the comparison and selection of human-centered design methods	Design process, duration	Conceptualized analysis and pre-defined dimensions
(Sanders et al. 2010)	Participatory design	To compare, discuss and deciding tools and techniques for participatory design	Form, purpose, context	Conceptualized analysis and pre-defined dimensions
(Vermeeren et al. 2010)	User experience	To provide an overview of the available UX methods	Development phase, participants, application, time length, data types, origin, location	Workshops
(Alves and Nunes 2013)	Service design	To guide newcomers and support communication in a team	What, when, where, who, why, how	Conceptualized analysis and pre-defined dimensions
(Curedale 2013)	Service design	To introduce methods to designers for designing services	Define the intention, framework, know people and content, explore ideas	-

In the field of design thinking, Roschuni et al. (2015) focus on supporting idea exchange as well as communication. By using several workshops, design methods are categorized into five main dimensions based on five design phases: research, analysis, ideate, build, communication. Under each dimension there are at least six characteristics (e.g., purpose, reflection time, participants); under each characteristic, there are sub-characteristics (e.g., using short-term and long-term to describe reflection time). With such a large number of characteristics, although a design method can be described in great detail, it is difficult for people to use them as a filter to select design methods (Miller 1956; Nickerson et al. 2013). Whereas the classification of design techniques in the field of service innovation is based on the design inputs, process, outputs, and benefits without detailed description on the approach for developing the classification (Kumar and LaConte 2012). In the research of design methods, the dimensions are more detailed by not only including the design process and purpose but also describing the content of the techniques and the participants (Martin and Hanington 2012). In the field of human-centered design, design methods are organized into the dimension of the design process and duration as bases for selecting appropriate techniques for specific design situations (Maguire 2001). As the number of methods and techniques expanded, further dimensions such as difficulty and the time cost of using the methods have been added based on case studies (IDEO 2015). When the design process emphasizes the participation of users (i.e., participatory design), techniques are categorized based on the dimensions form (action forms between participants), purpose (targets of using tools and techniques), and context (where and how tools and techniques are used). Under each dimension, three to four characteristics are available for people to allocate tools and techniques. The dimensions and characteristics are pre-defined based on experts' knowledge (Sanders et al. 2010). The classification of design techniques for UX evaluation includes seven categories (development phase, participants, application, time length, data types, origin, location) based on several workshops (Vermeeren et al. 2010). In the field of service design, techniques are classified differently. For example, with the purpose of introducing design methods, four classes (define the intention, use a framework, know people and content, explore ideas) are used (Curedale, 2013). While, for supporting the communication in a team and guiding novices to use appropriate tools and methods, Alves and Nunes (2013) build a taxonomy with six dimensions (what, when, where, who, why, how). The described classifications above serve a variety of different purposes that do not fit the digital service domain.

Table 2-3. Similarities and differences across the dimensions

Study	**Design phase**	**Participant**	**Time duration**	**Context**	**Data type**	**Evaluation type**	**Purpose**	**Origin**	**Content**	**Form**	**Benefit**	**Application**	**Difficulty**
(Roschuni, Kramer, Qian, et al. 2015)	x												
(Kumar and LaConte 2012)	x			x							x		
(Martin and Hanington 2012)	x	x			x		x	x	x				
(IDEO 2015)	x	x	x										x
(Maguire 2001)	x		x										
(Sanders et al. 2010)				x			x			x			
(Vermeeren et al. 2010)	x	x	x	x	x			x				x	
(Alves and Nunes 2013)	x	x		x		x	x		x				
(Curedale 2013)	x						x		x	x			

Note: in this summary, only the main categories are included, the sub-categories and the sub-sub-sub-categories are not included

In the above-introduced classifications, some similarities and differences across the categories can be recognized (Table 2-3). Except for Sanders et al. (2010), eight studies mention design process as a dimension; four studies mention purpose as a category in the classification of design techniques (Alves and Nunes 2013; Curedale 2013; Martin and Hanington 2012; Sanders et al. 2010); four studies emphasize the participation of users (Alves and Nunes 2013; IDEO 2015; Martin and Hanington 2012; Vermeeren et al. 2010); three studies describe time duration as an independent dimension (IDEO 2015; Maguire 2001; Vermeeren et al. 2010), which is only mentioned as an issue to be considered but not as a dimension or characteristic in Sanders et al. (2010). Except for Maguire (2001), none of the studies focuses explicitly on the goal of helping with the selection of design techniques. As the number of design techniques is growing, a better understanding of the different dimensions for classification is required. Besides, in previous research, a detailed explanation of the underlying research approach for developing and evaluating the proposed taxonomies is missing. Three of the studies do not mention the underlying research approaches (Curedale 2013; Kumar and LaConte 2012; Martin and Hanington 2012). Another weakness is the confusion of the three key concepts introduced above: design methods, design techniques, and design tools. One of the studies classifies design methods and tools together into a taxonomy (Alves and Nunes 2013). However, the confusion of the concepts causes overlaps in the dimensions. For example, it is difficult to classify the method usability test into different design phases because usability testing includes not only procedures, but also a set of principles that can be used in many design phases. Thus, overlaps will appear between dimensions, which are generally viewed as a weakness of any taxonomies (Bailey 1994; Nickerson et al. 2013). Although in practice, it can be helpful to combine methods, techniques, and tools for a specific design situation (Woolrych et al. 2011), it is necessary to have a clear overview of design techniques before going a step further to develop a systematic approach for the design process. Although the idea of classifying design techniques to show the similarities and differences in between is not completely new, it is still important to build a concise classification of digital service design techniques following a rigorous research approach.

2.6.2 Publicly available Web-based Platforms

By searching the Internet with the keywords: service design tools/techniques, usability techniques, UX methods, or design methods, one can easily identify several web platforms. The main objectives and features of these tools are summarized in Table 2-4. All of the web-based platforms have the objective of providing information about design tools, techniques, methods, and methodologies. Each of these tools provides several categories for supporting the selection process. However, there are several ambiguities in the way this is currently done.

Table 2-4. Related existing web-based platforms

URL	Objectives	Features
usability.gov	• Provides overviews of the user-centered design process and other user-experience disciplines • Covers methodologies and tools for making usable and useful content	• Selection based on eight categories • Detail descriptions of each design method
servicedesigntools.org	• An open knowledge platform (methodologies and tools) for the design research community	• Selection function based on four main categories with sub-categories • Short descriptions of each tool or method
allaboutux.org	• A shared information platform for collecting experience from the user experience community	• Selection based on four main categories with sub-categories • Short description of each evaluation method • Comment function
medialabamsterdam.com/toolkit	• Provides a collection of design methods • To help design participants to plan and execute design research within short iterations	• Filter function based on four parameters • Short description and instruction of each method
thedesignexchange.org	• To help design participants find potential methods for any problems • Developing a community of design educators and practitioners to evaluate the methods and educate the next generation of design innovators	• Selection function based on five categories with so many sub-categories and sub-sub-categories • Detailed or short descriptions • Create new methods after login

First, the unclear definition of the terms “tool, technique, method” is a reason that may cause confusions when trying to help design novices to select suitable techniques. The definitions of these terms present that design methods are very generic and need systematic thought, whereas, design tools are specific instruments that may change and may have many different versions. As design techniques provide steps in a procedure, design techniques can be classified into categories for supporting the selection (Brinkkemper 1996; Sanders et al. 2010). Thus, it is necessary for a web platform with the purpose of helping design novices focusing on design techniques.

Second, plenty of categories and sub-categories can be another reason that causes confusions. For example, the web platform thedesignexchange.org explicitly articulates an objective of helping design participants to choose design techniques by providing five main categories; under each main category, there are at least six sub-categories; under each sub-category, there are more than three sub-sub-categories. However, based on the studies from Nickerson et al. (2013) and Miller

(1956) who argue that the number of categories should not be overwhelming, the proposed classification is too complex (Miller 1956; Nickerson et al. 2013). For the purpose of supporting the selection process of design techniques, it is necessary to include a classification with a limited number of categories.

Third, when referring to enabling experts and design novices to exchange design knowledge, the existing tools that provide a similar function are allaboutux.org and thedesignexchange.org. Allaboutux.org provides a comment function, which enables people to give feedback and ask questions. But there is no feature of co-creation of classification. In contrast, users of thedesignexchange.org can suggest new design techniques, which helps to improve the content on the web platform. But the users are enabled to use the pre-defined categories to suggest the classification. Maybe there is no pre-defined category that is consistent with the categories in the users' mind. Hence, besides leaving comments and feedback, the users of the web platform should be enabled to bottom-up suggest categories in order to better support the selection of design techniques.

Based on the analysis above, although each of the existing web platforms has specialties on different purposes, there are common drawbacks when helping design novices in the selection of design techniques, such as unclearly defined terms, lack of consideration of the bottom-up suggested categories.

3 Research Methodology[1]

In order to answer the main research question and each break-down research question, this research project followed the DSR methodology (Vaishnavi and Kuechler 2004). As the purpose was to help novices select design techniques, not only an artifact for decision support needs to be provided, but also the effect of the artifact on the selection needs to be analyzed. Thus, I also combined design science research and behavioral science research. The two research paradigms complement each other (Hevner et al. 2004).

Figure 3-1 presents three design cycles, and each cycle was to answer a break-down research question. There were five steps in each cycle. After identifying problems, potential solutions were suggested, and developments were conducted, then the developed artifacts were evaluated.

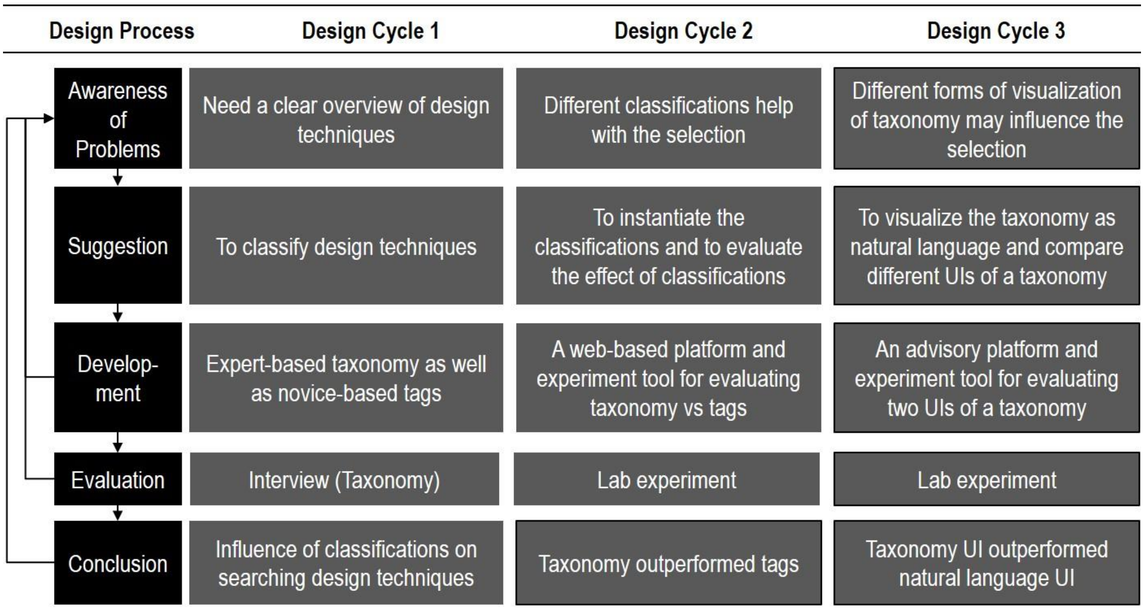


Figure 3-1. Design cycles by using DSR methodology, adapted from Vaishnavi and Kuechler (2004)

Design Cycle 1

In cycle 1, the problem that the expansion of the number of design techniques increases the difficulties when selecting appropriate ones was identified (Bollen et al. 2010; Greifeneder et al. 2010). A suggestion to this problem is to develop a tool for supporting the selection process. But before developing an artifact to help with the selection of design techniques, a comprehensive

[1] This Chapter is based on Liu (2018).

overview of the existing design techniques needs to be provided. Thus, building classifications that present the similarities and differences between design techniques can be a starting point (Bailey 1994). As this research seeks to help novices select design techniques, and there are differences between novices and experts when solving problems (Schenk et al. 1998), classifications from both experts' and novices' perspectives need to be considered. Hence, two classifications were developed. One was based on experts' knowledge (i.e., a hierarchical taxonomy), the other reflected novices' understandings of the content of design techniques (i.e., tags). The taxonomy was developed by following Nickerson et al. (2013), which resulted in a taxonomy of design techniques with five dimensions and 15 mutually exclusive and collectively exhaustive characteristics. The tags were created by following Spencer (2009) and Rugg and Mcgeorge (1997) and resulted in 16 unstructured categories. The taxonomy was evaluated by interviewing design experts (Kitchenham et al. 2005; Schultze and Avital 2011), and the differences between the taxonomy and tags were compared. The conclusion of cycle 1 discussed that the classifications needed to be instantiated as a tool to help with the selection, which was used as the basis for cycle 2.

Design Cycle 2

In cycle 2, the created classifications were instantiated in a web-based platform. Design principles were provided for the development of the platform for supporting the selection of design techniques. The usability of the platform was evaluated with a pilot filed evaluation (Duchowski 2002; van den Haak et al. 2006; Nielsen and Pernice 2009; Wixom and Todd 2005). However, the usability test could not provide detailed insights into the performance of the differences between taxonomy and tags when design novices select design techniques. Thus, in order to evaluate the effect of different types of classifications (taxonomy vs. tags) on the novices' selection of design techniques in-depth, a between-subject lab experiment was conducted. Hypotheses and a research model were proposed based on the cognitive fit theory with the variable of selection accuracy, cognitive effort, and rational as well as intuitive decision-making styles (Engin and Vetschera 2017; Hamilton et al. 2016; Vessey 1991). The result showed that the hierarchical taxonomy outperformed the tags in the task of selecting design techniques. In cycle 2, it was visualized as a graphical UI with abstract dimensions and characterisitcs of design techniques. But there are other ways to visualize a taxonomy. For example, a natural language dialog embedded with the taxonomy can be developed (Chen et al. 2017). Thus, the web-based platform in cycle 3 was iterated with the instantiation of two UIs (i.e., taxonomy and natural language dialog). The influence of these two UIs on selection of design techniques was analyzed in cycle 3.

Design Cycle 3

In cycle 3, new design principles for presenting the dimensions and characteristics of the taxonomy were provided (Arora et al. 2013; McTear 2002). The design feature that addressed the new design principle was implemented (Meth et al. 2015). Specifically, a chat frame with a dialog system that based on finite state-based system and frame-based system to suggest design techniques was proposed (Allison 2012; Arora et al. 2013; McTear 2002; Di Prospero et al. 2017). Two forms of UIs (taxonomy vs. natural language dialog) were evaluated and compared by using a within- and between-subject lab experiment. The effect of different UIs on novices' selection of design techniques with the consideration of rational decision-making styles was analyzed (Hamilton et al. 2016). The result presented that overall, the taxonomy UI outperformed the natural language UI in selection accuracy and perceived task performance. An in-depth analysis reflected that the taxonomy UI was more appropriate for high rational decision-makers and long-term use, whereas the natural language UI was more suitable for low rational decision-makers and short-term use.

The development and evaluation process in the three design cycles are introduced separately in detail in Chapter 4 (cycle 1), Chapter 5 (cycle 2), and Chapter 6 (cycle 3).

4 Design Cycle 1: Classifications of Design Techniques[1]

In the first design cycle, artifact development was separated into two parts. Part 1 describes the top-down development of the taxonomy from the perspective of design experts, while part 2 describes the bottom-up development of tags from the perspective of design novices. Figure 4-1 presents an overview of the two parts. Details are explained in the following.

	Part 1 (Chapter 4.1)	Part 2 (Chapter 4.2)
Classifications	A taxonomy of design techniques	Tags of design techniques
Experts or novices	Design expert knowledge based	Design novice knowledge based
Main method	Taxonomy development method	Card sorting
Key deliverables	Five dimensions and 15 mutually exclusive and collectively exhaustive characteristics	Not hierarchically structured 16 categories (i.e., tags)

Figure 4-1. Research approach for design cycle 1

4.1 Expert-based Taxonomy

In this Chapter, the development, evaluation, and conceptualization of the expert-based taxonomy are explained.

4.1.1 Research Approach

The research approach for developing the taxonomy was divided into three stages as depicted in Figure 4-2. In stage 1, an initial version of the taxonomy was created by following Nickerson et al. (2013)'s taxonomy development method. In stage 2, Kitchenham et al. (2005)'s framework and expert interviews (Schultze and Avital 2011) were applied to evaluate the content of the initial version of the taxonomy (i.e., dimensions and characteristics). Based on the evaluation, wording and the conceptualization of each dimension and characteristic in the initial version were optimized in stage 3.

[1] This Chapter is based on the following studies which are published or in work: Liu et al. (2016); Liu, Werder, et al. (2019a).

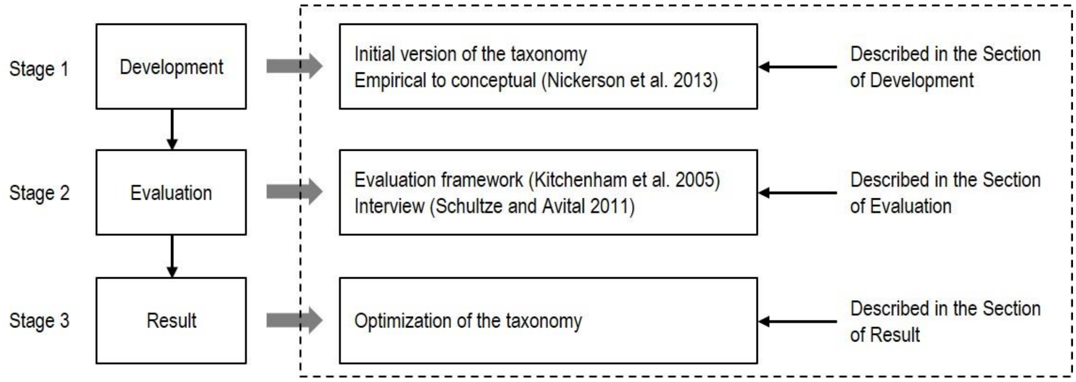


Figure 4-2. Research approach for developing the taxonomy of design techniques

4.1.2 Development

A taxonomy is a set of dimensions, and each dimension consists of a set of two or more mutually exclusive and collectively exhaustive characteristics so that each object (i.e., design technique) can be classified into one and only one characteristic for each dimension (Nickerson et al. 2013). Characteristics are defined as features that reflect the similarities and differences between the classified objects (Nickerson et al. 2013). Taxonomies have been widely applied in many fields, e.g., in biology and medicine. In addition, taxonomies typically reflect the structure and hierarchy of a website, which helps users to find the information they need (Boulton et al. 2016). Hence, a taxonomy that provides a clear overview of the broad spectrum of available design techniques as a starting point is needed to support the selection of design techniques.

Nickerson et al. (2013)'s method was followed to develop a taxonomy of design techniques, which provides a detailed description of the principles and steps for developing taxonomies. The specific method was chosen for four reasons. First, this method suggests developing a taxonomy by identifying a subset of objects (i.e., design techniques). Given the fact that new design techniques appear all the time, it is challenging to obtain a list of all design techniques available. Therefore, the taxonomy can be developed in an iterative manner, which extends the taxonomy as new data sources. Second, the method provides subjective and objective ending conditions. These ending conditions not only guide the development process by preventing endless iterations during the development, but also suggest when further evolutions of the taxonomy are needed. When new design techniques are introduced, which cannot be assigned to any of the existing dimensions or characteristics, new dimensions or characteristics can be added to extend the taxonomy. Third, by following this method, mutually exclusive and collectively exhaustive characteristics will be created, which give a clear overview of design techniques. Fourth, the method has been successfully applied; for example, developing a taxonomy of evaluation methods for information systems artifacts (Prat et al. 2015), and developing a taxonomy to analyze the IT value literature. To summarize, Nickerson et al. (2013)'s taxonomy development method was considered as suitable to develop the taxonomy of digital service design techniques.

Taxonomy Development Process

Figure 4-3 depicts the steps that were followed to develop the taxonomy of design techniques.

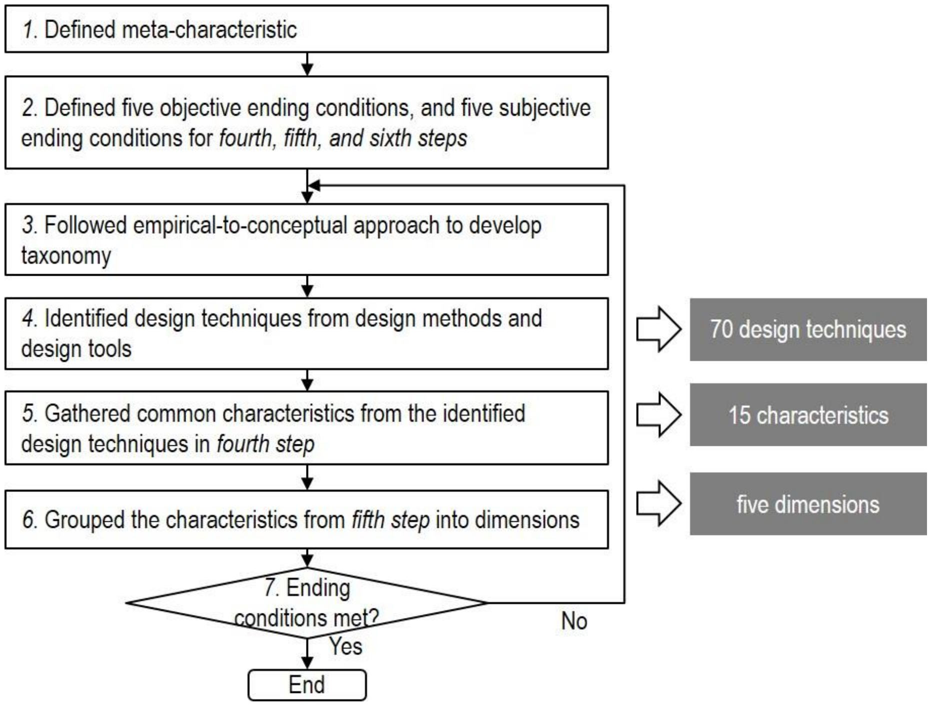


Figure 4-3. Taxonomy development steps

In the *first step*, the meta-characteristic and purpose of the taxonomy were determined, i.e., to develop a taxonomy with mutually exclusive and collectively exhaustive dimensions and characteristics that provide a comprehensive view of design techniques and to serve as a foundation for future research, supporting the selection of design techniques.

In the *second step*, the ending conditions that were used to terminate the iteration in the taxonomy development procedure were defined. Two types of ending conditions - objective ending conditions and subjective ending conditions - were formulated (Nickerson et al. 2013). Objective ending conditions tested whether each dimension had mutually exclusive and collectively exhaustive characteristics. Subjective ending conditions checked whether it was a useful taxonomy from a theoretical perspective. In order to test the taxonomy with these ending conditions systematically and purposefully, the applicable ending conditions were organized in Table 4-1, presenting the ending conditions for each step (cf. Prat et al. 2015). The detailed steps included identifying design techniques (fourth step), identifying characteristics (fifth step), and grouping characteristics into dimensions (sixth step). To ensure that the judgment of the subjective ending conditions was unbiased, two people that were involved in the development of the taxonomy evaluated the subjective ending conditions together.

Table 4-1. Ending conditions for developing the taxonomy of design techniques

Ending conditions	Steps
Objective ending conditions	
All design techniques have been examined and cannot be merged or split.	4
At least one design technique is classified under each characteristic.	4, 5
Each characteristic is unique and cannot be repeated (no characteristic duplication).	5
Each dimension is unique and cannot be repeated (no dimension duplication).	6
No new dimensions or characteristics can be added in the last iteration.	5, 6
Subjective ending conditions	
Concise: the number of dimensions is not un-widely or overwhelming	6
Robust: enough dimensions and characteristics to classify design techniques	5, 6
Comprehensive: all design techniques should be classified within the taxonomy	4, 5, 6
Extendible: a new design technique, characteristic, and dimension can be easily added	4, 5, 6
Explanatory: the dimensions and characteristics can explain design techniques	5, 6

In the *third step*, the empirical-to-conceptual approach was followed to develop the taxonomy based on a systematic and detailed analysis of secondary data (Myers 2009). There are sources with plenty of design techniques and explanations for the usage of each design technique.

In the *fourth step*, design techniques were distinguished from design tools and methods because of the objective to develop a taxonomy of digital service design techniques. In the design process of digital services, the service ecologies, UX, digital transactions, and variety of activities (e.g., generating ideas, creating and evaluating prototypes) need to be considered (Kutsikos et al. 2014; Mager 2009; Williams et al. 2008). The used sources should contain techniques that fit the context of digital service design. Thus, two people jointly decided the ending conditions that were used in the sources[1] from the academic and grey literature in order to choose related design techniques, which focus specifically on UI design, user-centered design, usability, and UX design. The

[1] i) *usability.gov* is a leading resource for UX practices and introduces 32 design and evaluation methods; ii) *servicedesigntools.org* is a research project conducted Research & Consulting Center of Domus Academy, which contains 36 methods, techniques and tools for service design; iii) *allaboutux.org* is the result of a survey conducted by Vermeeren et al. (2010) that provides a list of 82 evaluation methods for UX; iv) *Service Design* is a book that includes 250 methods, techniques, and tools for service desing (Curedale 2013); v) *Universal Methods of Design* is a book that includes 100 methods and techniques for widespread use in product and service design (Martin and Hanington 2012).

filtering of all repeated data points results in a list of 292 design methods, design techniques, and design tools. As design methods, design techniques, and design tools were often mixed up, design techniques were distinguished from design methods (e.g., crowdsourcing, experience sampling method, evidence-based design) and design tools (e.g., AttrakDiff, AEIOU, Attrak-Work questionnaire) based on the descriptions of methods, techniques, and tools. The final list consisted of 82 design techniques. Thereafter, design techniques with similar definitions were merged. First, the technique with a more precise name that well-explained the technique was chosen (e.g., concurrent think-aloud and retrospective think-aloud were chosen instead of think-aloud). Second, if neither name was considered more precise, the design techniques that had been used most often was selected for the final list (e.g., affinity diagramming was chosen instead of KJ technique). In sum, the number of design techniques decreased to a list of 70 design techniques.

In the *fifth step*, all the characteristics of the 70 design techniques were gathered. Since not all characteristics were directly presented in the explanation of the sources, the descriptions of each design technique into different characteristics were summarized by applying open coding (Corbin and Strauss 2014). The process of open coding was exploratory and can be used to identify concepts. In the coding process, the first coder coded the content, while the second coder checked the coding results. The two coders were the two researchers who evaluated the ending conditions and chose design techniques. For example, if the technique was described that it can be used in a short time period or a couple of days, code “short-term” was given. Table 4-2 presents some examples.

Table 4-2. Illustration of characteristics

Code	**Characteristics**
Short-term	Episode, momentary, short test task, short time period, minutes, hours
Long-term	Long time period, days, weeks, months, over time study, time-consuming
Real-time	Seconds, immediate, video record, observation, concurrent

The *sixth step* was to group all the characteristics from the fifth step into dimensions. In this step, all the characteristics were summarized into five dimensions. Besides the dimensions that were included in the prior studies, such as design phases, duration, participants, and evaluation types, time dependency (real-time and retrospective) was added as a new dimension. Because real-time and retrospective are two essential characteristics in design processes when identifying design techniques. While real-time describes users’ immediate affective response, retrospective is an indicator of people’s cognitive experience, which can be used to describe a preceding episode of use that happens immediately in the early phase of user test, or a period of days or weeks ago (Hassenzahl and Ullrich 2007; Kim et al. 2008). Thus, time dependency needs to be considered as a dimension.

In the *seventh step*, the derived dimensions and characteristics in the taxonomy finally met the criteria of the ending conditions (Table 4-1). A taxonomy with mutually exclusive and collectively exhaustive dimensions and characteristics was proposed (Appendix A).

4.1.3 Evaluation

An overview of the taxonomy evaluation is summarized in Figure 4-4. In the evaluation, experts with domain knowledge who had not involved in the development of the taxonomy participated in a face-to-face interviews. To structure the interview, an existing framework (Kitchenham et al. 2005) and appreciative and laddering methods (Schultze and Avital 2011) were used. The evaluation was based on the considered design situations in real design projects. The ínterviews were analyzed both qualitatively and quantitatively. In the following, the general evaluation method and the data collection, data analysis, and the results are explained.

WHO? Subject of evaluation	Involvement in taxonomy building	Experts participated in the evaluation were not involved in the taxonomy development
	Background	Participants came from both academics and practitioners.
	Experience	Experts involved in the evaluation worked in the design related domains (e.g., UX, HCI, product design)
WHAT? Object of evaluation	Type	The evaluation was based on the design situations that the experts considered in real design projects.
	Involvement in taxonomy building	Some of the design techniques used in the taxonomy building were mentioned by experts when describing design projects.
	Coverage	The described design techniques were selected by the experts when describing the design projects.
HOW? Method of evaluation	Approach	Qualitative: closed and open coding of the interview transcripts (Myers 2009) Quantitative: agreement test (Krippendorff 2004) and precision test (Fawcett 2006)
	Method	Expert interview (Kitchenham et al. 2005; Schultze and Avital 2011)

Note: The framework used for summarizing the evaluation approach is adapted from Szopinski et al. (2019).

Figure 4-4. An overview of the taxonomy evaluation

Taxonomy Evaluation Method

Interviews generate uplifting, affirmative experience (Schall et al. 2004) and therefore, 15 experts[1] were interviewed in order to evaluate whether the content of the taxonomy is consistent with the experts' experiences. Expert interviews provide a source to evaluate whether the content of the taxonomy is understandable and consistent with experts' working experience (Schultze and Avital 2011). Based on the evaluation purpose, an interview protocol with appreciative questions and laddering questions, the proposed taxonomy, and a questionnaire of demographic information was prepared. All questions were asked in a positive way in order to generate rich data without limiting the experts' thoughts. An adapted evaluation framework was used (Table 4-3) to structure the interview questions (Kitchenham et al. 2005).

Table 4-3. Adapted framework for evaluating the taxonomy based on Kitchenham et al. (2005)

Aspect	Description	Activity
Semantic quality	The taxonomy is complete and valid	Evaluate whether there are no new dimensions or characteristics suggested by the experts in the interviews
Test quality	The identified dimensions and characteristics in the taxonomy are considered in the design process	Evaluate whether the dimensions and characteristics are mentioned by the experts during interviews
Syntactic quality	The description of the dimensions and characteristics in the taxonomy are syntactically correct	Evaluate whether the description of the dimensions and characteristics are consistent with experts' descriptions
Pragmatic quality	The taxonomy is presented in an understandable format	Evaluate whether the proposed taxonomy can be understood by the experts in interviews

The interviews consisted of two parts (Appendix B). The first part aimed to evaluate semantic quality, test quality, and syntactic quality of the taxonomy. Experts were asked to describe the design process and selected design techniques in different design situations when designing digital services. In the second part of the interview, the initial version of the taxonomy was presented to collect experts' feedback in order to evaluate the pragmatic quality of the taxonomy.

Data Collection

All interviews were conducted in English and in-person with participants of a large academic- and practice-oriented human-computer interaction (HCI) conference. The conference provided a good opportunity to have direct access to design experts. Experienced participants from industry

[1] 15 experts were considered to be sufficient to evaluate the proposed taxonomy (cf. Dell and Kumar 2016; Vigo et al. 2014).

and academia were recruited. Before starting the interview, experts' expertise, working domains, and experiences were asked in order to identify suitable candidates for our interviews. As digital service design processes usually need the collaboration of people in the interdisciplinary domain (e.g., interaction design, software engineering), the interviewed experts were from a broad field (Table 4-4). Interviews took place face-to-face and were audio recorded after requesting permission from experts. During each interview, additional notes were taken. Each interview took 20-45 minutes. All interview recordings were transcribed.

Table 4-4. Participant profiles

Pseud-onym	Industry / Academia	Experie-nce years	Age	Education Background	Working domain
Alice	Industry	1-5	31-40	Communication psychology	Requirement engineering
Bert	Industry	6-10	41-50	Interaction design	Interaction and product design
Carl	Industry	6-10	31-40	Computer science	Interaction and product design
David	Industry	11-15	31-40	Software development	Product design
Emma	Industry	6-10	31-40	HCI	HCI
Frank	Academia	20-25	50-60	Information systems	Design oriented research
Grace	Academia	6-10	50-60	Usability	User-support programming
Harry	Industry	1-5	21-30	Business psychology	UX research
Irene	Academia	6-10	31-40	Computer science, software engineering	Research
Jan	Academia	11-15	50-60	Software engineering, mathematics	Software engineering, teaching
Kevin	Academia	6-10	41-50	Software engineering	Research, programming
Lisa	Academia	6-10	41-50	Computer science	Design oriented research
Max	Industry	1-5	31-40	Communication design, UX	Service design
Neil	Industry	6-10	31-40	Software engineering	UX design, usability engineering
Oliver	Industry	6-10	50-60	Graphic design	Product design

Data Analysis

In the analysis process, closed coding and open coding were used to evaluate the taxonomy (Myers 2009). QDA Miner 5[1], a qualitative data analysis software, was used to analyze the transcripts. First, the interviewer and a second coder who was unfamiliar with the interviews and

[1] https://provalisresearch.com/products/qualitative-data-analysis-software/

the taxonomy development process conducted closed coding independently (Myers 2009); the codes were the characteristics and dimensions in the taxonomy Krippendorff's alpha was used to measure the agreement between two coders because of the small sample size of experts (Krippendorff 2004). The percentage of the code co-occurrence was 89.8%, and Krippendorff's alpha was 0.765 (Cohen's kappa: 0.764; Scott's pi: 0.765), which was considered as acceptable reliability (Krippendorff 2004). The coding frequency results presented that all dimensions and characteristics in the proposed taxonomy were considered to be relevant in the design process (Appendix C).

Second, open coding was conducted by the interviewer to analyze the precision of the proposed taxonomy (Fawcett 2006). Precision means the proportion of positive results in statistics that are true positive results, which is often used in medical diagnoses (Fawcett 2006). In this research, the following formula was applied to calculate precision:

$$Precision = \Sigma\ True\ positives / (\Sigma\ True\ positives + \Sigma\ False\ positives). \quad (1)$$

Where *true positives* represent the number of confirmed characteristics from the interviews and *false positives* represent the number of additionally identified characteristics, *precision* means the proportion of the mentioned times of characteristics during the interviews were included in the proposed taxonomy. The large precision (93.78%) indicated that most of the characteristics in the proposed taxonomy were mentioned by experts during interviews.

4.1.4 Results

The coding results suggested that the proposed taxonomy was consistent with the experts' experiences and presented understandable dimensions and characteristics. In the following, the adjustments and improvements made to the taxonomy and the test of the precision of the taxonomy are explained.

Adjustments to the Taxonomy

Experts suggested renaming the dimension participants (user involved, without user) in the initial version to user participation (user attendance, user absence). The interviews reflected that user involvement was emphasized by all experts. However, real users were not always available. Designers had to use fictitious users to analyze designs. The experts argued that during design processes, they always kept users in mind, even though real users were not included in design activities. For example, Irene told us, *"At this point, we don't think we work with users, because we don't always have users when we want. [...] We use storyboards to know what kind of* *situations and why users use our software."* Hence, user attendance and user absence were considered to be more suitable than the user involved, without user.

Draft prototyping and detail prototyping in the initial version were replaced by low-fidelity prototyping and high-fidelity prototyping. When design experts talked about creating prototypes, they used the terms high-fidelity prototype or mock-ups to describe a functional prototype. As David shared, *"Next, we build maybe high-fidelity prototyping in the very early stage to get a better view and a better usage of the functions or some features."* In addition, Max told us, *"So the company suggested to create mock-ups, which could be used in the smartphone."* As mock-ups can be a pilot version with functions, the term high-fidelity prototyping was used. When mentioning draft prototyping, experts used paper prototypes and sketches to describe what artifacts they created in this design phase. For example, David stated, *"Next, we create some early sketches in an App. [...] The sketches are often used in our customer meetings." "Normally, I always begin with pen and paper. I don't know [...] just sketching. After the sketches, we choose which interface and make another one."* mentioned Irene. As paper prototypes and sketches may not be appropriate to describe a design phase, the term low-fidelity prototyping was used, which not only echoed high-fidelity prototyping but also simplified the understanding.

Based on the experts' descriptions, the term release was used instead of launching in the initial version for two reasons. On the one hand, experts tended to use the term release when describing a functional product. As David reported, *"At the end, we release the App and try to follow some approaches, which directly deliver some feedback which is really close to us we can really get more information over a long period of time."* On the other hand, launching implies that a product moves toward the market, while a release includes a pilot test where user data is collected before product launching. As stated by Frank, *"Then we typically start the pilot. That means they can use real work practices and gave us feedback on what they would like to have different."*

Precision of the Taxonomy

In order to measure the agreement and disagreement between the characteristics in the proposed taxonomy and the characteristics mentioned by the experts, the responses by the experts were matched with the proposed taxonomy based on the open coding (Table 4-5). As different people can create different taxonomies (Nickerson et al. 2013), the characteristics mentioned by only a few experts can be identified as potential extensions of the current version of the taxonomy.

Table 4-5. The characteristics mentioned by the experts that matched the proposed taxonomy

Class	Confirmed characteristics / True positives (number of observations)	New characteristics / False positives (number of observations)
Dimensions: Characteristics	Design phases: Planning (15) Low-fidelity prototyping (14) High-fidelity prototyping (14) Release (5)	Project types: Complex project (2) Detail project (1) Small project (1) New project (1)
	User participation: User attendance (14) User absence (11)	Collaboration: Stakeholders (1) Other departments (1)
	Duration: Long-term study (4) Short-term study (11)	Field or lab: Field (1) Lab (2)
	Time dependency: Real-time (11) Retrospective (7)	
	Evaluation types: Questionnaire (4) Interview (11) Experiment (5) Observation (12) Group discussion (13)	
Total	**151**	**10**

Note: The numbers in the brackets mean the numbers of experts mentioned the characteristic.

A distinction mentioned by experts was the difference between design activities in a lab or in the field. Lab experiments allowed designers to control the environment when presenting early prototypes. For example, Frank shared, *"We move to experiments. […] We can control the behavior so that we can make sure the weakness of our software doesn't go through the practice."* And Emma told us, *"We have a customer lab where we invite users to test our prototypes or later the finished project."* Going into the field enabled the designers to actually know how users use their product in a real scenario. Frank said, *"Then we typically start the pilot. That means they can use real work practices and give us feedback on what they would like to have differently."* The descriptions presented that lab and field related to some characteristics in the current version of the taxonomy. For example, the characteristic "experiment" was subsumed under the dimension of "evaluation type." To some extent, "experiment" might be seen as a duplication to "lab." The field analysis related to collecting user feedback in the early release phase. Adding a new dimension that describes field or lab explicitly in the taxonomy was considered. However, given the limited evidence, no change was made to the dimension.

In the proposed taxonomy, the dimension of "user participation" described whether users need actually to participate when using a particular design technique. However, practitioners pointed out that participants had different roles. Some design projects involved end users. Alice said, *"I work with engineers (automotive engineers)."* Others may not get access to such end-users and hence, involve other stakeholders (such as customers). *"We do the retrospective test, but only with stakeholder but not with real users,"* reported Neil. In addition, some projects involved people from different functions, as Oliver mentioned, *"We have people from different backgrounds. I talked with marketing and different people."* Such users may have different roles when using the system, e.g., as captured by personas. Hence, different roles were involved in the design process. However, in order not to overload the taxonomy, the taxonomy focused on the users in general.

Two experts suggested that different project types influenced the selection of design techniques. The complexity of a project could influence the decision for or against a specific design technique as Emma reported, *"When we have a web application or a desktop application. They are very big and complex; then we use think-aloud. When we have little small phone Apps or focus only on one little feature, then we also try to use eye-tracking to get the last 20%."* Besides project complexity, participants also reported that the maturity of the project influenced the selection of a design technique. Carl shared, *"It's like kind of new software or real big new function then we start with brainstorming and workshop sessions. […] It's more like the basic information we have, like how big the feature is, and how new the software is."* This situation reflected that the choice of design techniques could be characterized by the attributes of projects. However, the descriptions presented that the maturity of a design project might relate to the design phase. For example, in the beginning, a project is broad and complex; when it moves to a prototype, the features become more specific. Hence, it was not clear enough to add design project as a dimension in the taxonomy.

4.1.5 Conceptualization and Exemplary Application of the Taxonomy

Based on the development and evaluation, the dimensions and characteristics of the final taxonomy are conceptualized. Furthermore, an exemplary application of the developed taxonomy is demonstrated.

Conceptualization of the Taxonomy

Figure 4-5 provides a compressed overview of the final version taxonomy for digital service design techniques.

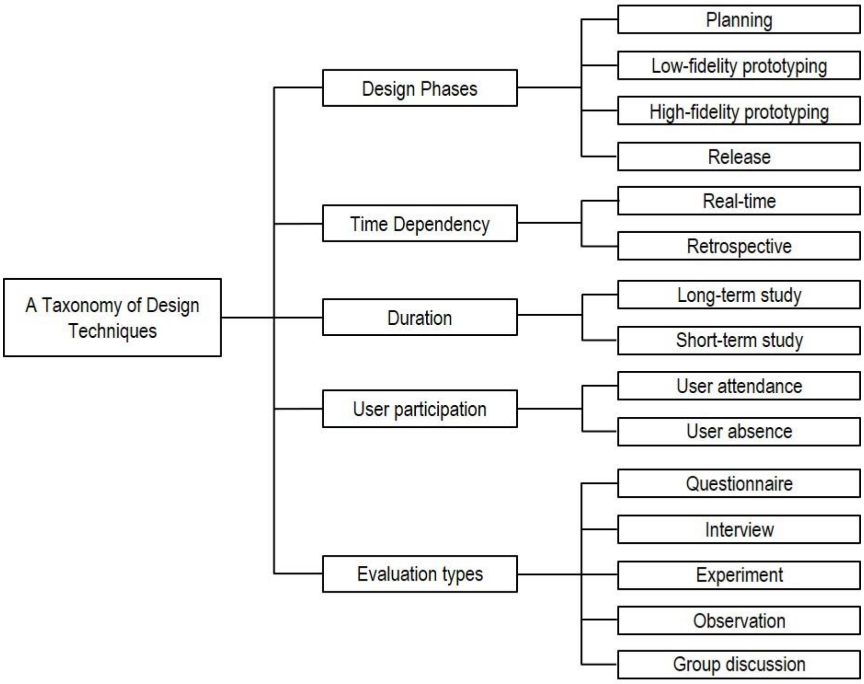


Figure 4-5. A taxonomy of digital service design techniques

Design phases (planning, low-fidelity prototyping, high-fidelity prototyping, release) are conducted in an iterative cycle within design processes. In the planning phase, design techniques are usually applied to analyze shared design ideas (e.g., using territory maps to visualize shared focuses of a design team), and to collect and analyze context and users (e.g., using touchpoint matrix to analyze the connected UIs in a digital service or applying shadowing to observe how users interact with a digital service). In the low-fidelity prototyping phase, designers use design techniques to create low-fidelity prototypes (e.g., applying collaborative sketching to create prototypes with collaboration) and to make a series of assessments to compare the prototypes (e.g., using the repertory grids to reveal which paper prototype satisfies users). In the high-fidelity prototyping phase, design techniques focus on the in-depth evaluation of design work and the modification of the prototype to get a stable version with basic functions, which can be used to analyze detailed features and to collect user feedback (e.g., using the concurrent think-aloud to detect usability problems). In the release phase, design techniques help to collect data from a pilot version or an internally released version or a publicly released version to evolve their designs, and to prepare for the next iteration (e.g., using web analytics to understand web usage better in order to improve the design).

Time dependency discriminates between immediate affective feedback and retrospective cognitive feedback based on memory. The collection and analysis of real-time feedback reveal real-time use and users' affective change when interacting with designs (e.g., using the eye-tracking technique to analyze users' reading habits). Retrospective feedback is collected based on the users' memory of their experience of using a digital service in a previous episode that happened immediately or early before (e.g., using product experience tracker to gather user feedback periodically).

Duration describes the time length of an iteration cycle of a design process. Long-term studies need to be conducted over a long period of time. They present how relationships between services and users evolve over time and identify trends in users' satisfaction (e.g., collecting users' day-to-day experiences by applying long-term diary study). Short-term studies can be used to conduct design activities or data analysis in a rather limited period of time (e.g., using bodystorming to quickly evaluate low-fidelity prototypes).

The dimension user participation is used to decide whether real users are involved. When real users become involved in the design process, designers can observe, analyze, and predict how well their designs fit users' expectations (e.g., involving users to create low-fidelity prototypes by using flexible modeling technique). However, on some occasions, real users should not be involved, or involvement is too costly. For example, in the planning phase, if designers directly ask users for advice, users may have no idea of what exactly they expect. In this case, using fictitious user profiles is more suitable than getting real users to participate in design processes (e.g., analyzing user behavior patterns by using personas).

As the design process of digital services is iterative, evaluations take place not only at the end of the design process but also run throughout. Evaluation types indicate how to evaluate an outcome of design activity. For example, designers use the design technique scenario to create narratives, exploring how people will interact with a design. The narratives help a design team when discussing a design version and design goal. When a design technique is used within a team to share ideas or to make a decision, this technique was assigned into group discussion (e.g., applying stakeholder walkthrough to evaluate low-fidelity prototypes). The other four characteristics of evaluation types are questionnaires, interviews, experiments, and observation. Questionnaires enable the evaluation of users' goal achievement, satisfaction, etc. (e.g., collecting users' emotional feedback by using the "3E" technique to create a questionnaire). Interviews generate rich data that reveal users' experiences and expectations (e.g., using photo elicitation interviewing to evoke conversation and recall users' experience). Experiments reflect user behavior in a specific environment by means of concrete data (e.g., testing whether users are willing to do things in a new way by using the Wizard of Oz technique). Observation not only includes observing how users use a digital service and how they perform in their daily life, but also observing whether digital UIs fulfill rules to achieve high usability (e.g., using fly-on-the-wall observation technique to observe users without interference).

Exemplary Application of the Taxonomy

Table 4-6 presents an example of how the taxonomy can act as a starting point to help with the selection by narrowing down the search of design techniques.

Table 4-6. An exemplary use of the taxonomy for filtering design techniques (Appendix D).

Dimensions and characteristics / Design techniques	**Design Phases**	Planning	Low-fidelity prototyping	High-fidelity prototyping	Release	**Time dependency**	Real-time	Retrospective	**Duration**	Long-term study	Short-term study	**User participation**	User attendance	User absence	**Evaluation types**	Questionnaire	Interview	Experiment	Observation	Group discussion
When choosing ***planning, real-time****:*																				
Contextual Laddering		x					x				x		x				x			
Fly-on-the-Wall Observation		x					x				x		x						x	
Graffiti Walls		x					x			x			x						x	
Role-Playing		x					x				x			x						x
Shadowing		x					x				x		x						x	
…																				
Total number		23	0	0	0		23	0		1	22		6	17		1	1	0	3	18
When choosing ***planning, real-time,*** *and* ***observation****:*																				
Fly-on-the-Wall Observation		x					x				x		x						x	
Graffiti Walls		x					x			x			x						x	
Shadowing		x					x				x		x						x	
Total number		3	0	0	0		3	0		1	2		3	0		0	0	0	3	0

Notes: "x" means this characteristic is selected to filter design techniques; "Total number" indicates how many design techniques in each column of characteristics are left after selecting specific characteristics.

When a design project is in the planning phase, and designers prefer real-time study, these two design situations can be seen as the characteristic planning from the dimension design phases and the characteristic real-time from the dimension time dependency. When planning and real-time are selected from the taxonomy, the techniques that are classified under other characteristics from the dimensions of design phases and time dependency will be filtered out. Thus, the number of relevant design techniques that are suitable for the design situation decreases to 23. When designers plan to conduct design activities with the participation of real users with considering the two previous situations (i.e., planning and real-time), another characteristic user attendance is selected. Subsequently, the techniques classified under the characteristic of user absence in the dimension of user participation will be filtered out. Thus, the number of appropriate design techniques for the design project will be further reduced. Finally, only three techniques remain (i.e., fly-on-the-wall observation, graffiti walls, shadowing). The final decision of design techniques that appropriate for the design project can be selected from these three techniques. The example shows that the taxonomy not only provides a comprehensive overview of existing design techniques but also helps in effectively filtering design techniques.

4.1.6 Conclusion

In the first part of the design cycle 1, an empirical-to-conceptual approach was followed to develop a taxonomy of digital service design techniques. The taxonomy was evaluated by conducting qualitative research in order to validate and further improve the usefulness and acceptance of the taxonomy. In addition, an exemplary application of the taxonomy was demonstrated.

The development of the taxonomy contributes to digital service design from different perspectives. The concise and extensible taxonomy with five dimensions and 15 characteristics presents an overview of the similarities and differences of digital service design techniques. It represents a type I theory for analysis (Gregor, 2006), which can support building more systematic knowledge about design techniques. The taxonomy can be used as a foundation for future research studies contributing relevant knowledge to design processes as well as design techniques. For example, the proposed taxonomy is an artifact that can be used as a basis to support the analysis of the decision-making process of proper design techniques under different design situations. From a more practical point of view, the interview result presents that the proposed dimensions and characteristics are highly relevant to design processes, which means that the dimensions and characteristics can help to organize design techniques in design processes. The potential use of the taxonomy can further benefit digital service design processes (Glass and Vessey 1995; Lutters et al. 2014; McKay et al. 2012).

4.2 Novice-based Tags

Complementing the top-down taxonomy from the perspective of design experts, the bottom-up classification of design novices was also considered. In this Chapter, the open card sorting procedure for creating tags with design novices is described. The two different types of classifications were also compared.

4.2.1 Research Approach

The overall research approach was separated into three stages (Figure 4-6). In the first stage (section 4.2.2), an open card sorting with novices was conducted, and the created categories were collected. In the second stage (section 4.2.3), the created categories from novices were analyzed and standardized categories were built. In the third stage (section 4.2.4), the bottom-up classification was compared with top-down expert-based classification.

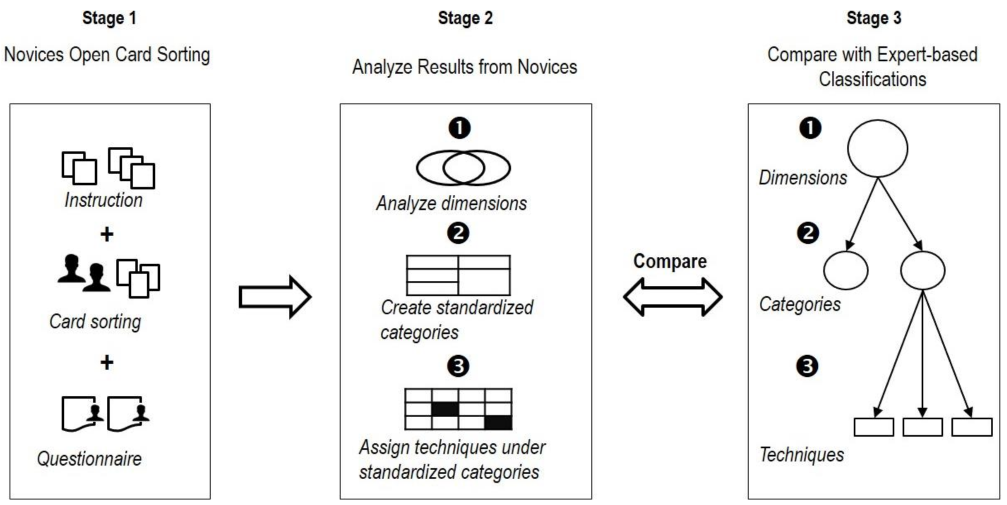


Figure 4-6. Research approach for creating tags

4.2.2 Open Card Sorting Procedure

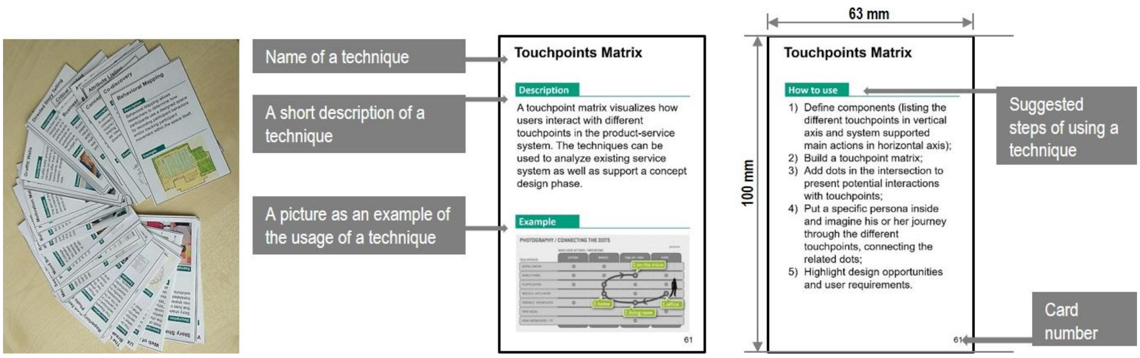
The repeated single-criterion sorts suggested by Rugg and Mcgeorge (1997) were followed in the sorting process. In repeated single-criterion sorts, participants were asked to sort the cards based on a single criterion (i.e., viewpoint) each time (Rugg and McGeorge 1997). As novices intend to sort entities from lower level categories (Chi et al. 1981), it is more flexible and easier for novices to handle when sorting techniques based on only one viewpoint (Rugg and McGeorge 1997). In the sorting process, participants were, at first, given instruction of card sorting. Subsequently, they started to sort design techniques into different categories and gave labels to the categories. Finally, they were asked to fill out a questionnaire as input for collecting their demographic information and the perceived confidence of the sorting result.

Participants

With a purpose of understanding how design novices classify design techniques, an open card sorting was conducted with 40 students studying computer science, information systems, industrial engineering and management with an average age of 22 years (SD = 2.2) (23 male, 17 female). The students had some basic knowledge of design, and they could be potential consumers of design techniques in their future job but did not have any real-world practical experiences when conducting the experiment. Thus, it was reasonable to consider students as novices and conduct open card sorting with them. In the open card sorting exercise, in order to avoid bias from a single person (cf., Burnay 2016), two students were randomly assigned to a team, and 20 teams were built. Two participants were asked to sort the cards together. Each team sorted cards without interruptions or influences from other teams.

Materials

As the card sorting process is suggested to include 30-100 cards (Spencer 2009), 70 self-explained small filing cards (i.e., 70 design techniques[1]) were used for sorting. Several envelopes were provided for collecting created categories. The suggestion of Rugg and McGeorge (1997) were followed to prepare small self-explained filing cards (Figure 4-7).



Note: The content and picture on the card were from Brugnoli (2011).

Figure 4-7. An example of a self-explained card

The size of each card was 63mm x 100mm to make sure the cards were easy to handle. Each card included five main information pieces. The front side included the name of the design technique,

[1] The 70 design techniques are selected based on the definition of methods, techniques, and tools (Brinkkemper 1996; Hackathorn and Karimi 1988; Kettinger et al. 1997) from the following sources: allaboutux.org (includes the methods in Vermeeren et al. (2010)); servicedesignkit.org (an application of Liu et al. (2016.)); introduced techniques in Martin and Hanington (2012). These three studies were used to compare novice- and expert-based classifications.

a short description of the design technique, and an example or a scenario of the design technique. The back side described a brief instruction on how to use the design technique. Both sides had a number of the card for the convenience of data collection. In order to collect the sorted cards and the created categories, participants were asked to put the cards that belong to the same category into an envelope and label each category and/or give an explanation to each category. The viewpoint that they based on for categorizing and the team number also needed to be written on the envelopes.

Sorting Steps

The open card sorting process was divided into three steps: reading a short instruction of card sorting, sorting cards into categories, and answering a questionnaire (Rugg and McGeorge 1997). *First,* what open card sorting is and how to sort cards was explained. In order not to bias participants, sorting examples from zoology was provided in the instruction. For instance, polar bears and koala bears can be classified as Mammalia based on the viewpoints of morphology. However, polar bear and koala bear belong to two categories, carnivore and herbivore, based on the viewpoint of eating patterns. Such examples were used to explain that the participants can create the design techniques based on different viewpoints and there is no wrong classification. *Second,* each team read the information on each card and sorted cards into categories based on their understandings of design techniques. The participants were expected to sort design techniques based on only one viewpoint and create no less than three and no more than nine categories (Nickerson et al. 2013; Rugg and McGeorge 1997). Each team put sorted cards into envelopes and wrote a label and/or a short description for each category on each envelope. If some cards could not be sorted into a category, students could put them into an envelope called "not applicable." *Third,* after each team finished the sorting, each participant was asked to fill out a questionnaire independently. The questionnaire was used for collecting demographic information, experiences of design projects, and the confidence level of the card sorting result from the sorting session. All participants were very confident about their sorting results. The open card sorting session for each team lasted between 45 to 60 minutes.

4.2.3 Results

After collecting the outcomes from open card sorting, the exploratory analysis approach was followed to analyze the result (Spencer 2009). As presented in Figure 4-6, there were three steps. *First,* the viewpoints that novices used as the basis for categorizing design techniques were analyzed and aggregated. *Second,* categories from the originally created categories with similar meanings were standardized. In the open card sorting process, there were no pre-defined categories. All the categories were created by the participants themselves. As 20 teams participated in the sorting exercise, 20 groups of categories were created. Different teams might create the same categories or use different labels for the same categories. Before comparing categories created by novices with the categories in the expert-based classifications, the same categories, and categories with similar meanings were merged to create standardized categories (Righi et al. 2013; Spencer 2009). *Third,* techniques were assigned into standardized categories. In the collected result, besides the created categories, the classified techniques under each category were also presented. An overview of the assigned design techniques into different categories was presented. In order to have a deep understanding of the overlap between design techniques and standardized categories, the percentage of the teams that agreed to the assignment of each design technique in the standardized categories were computed, which depicted the overall agreement of the allocation of design techniques from novices.

Viewpoints

The open card sorting result showed that 20 teams created 110 categories (excluding the category "not applicable") based on 11 different viewpoints. The used viewpoints were summarized in Table 4-7. More than 63.6% of the participants sorted cards by combining two or more viewpoints, even though students were told to sort the cards based on only one viewpoint. Take team 5 as an example; team 5 created five categories which include "use diagrams to visualize results," "only designer participation," "pre-experience scenarios to solve problems," "collaboration between user and designer," and "user-oriented". These five categories described two viewpoints: participants and activity types. For most of the participants, the combined viewpoint might refer to a unique viewpoint. The most mentioned viewpoint was purpose which was considered as a single viewpoint by three teams and appeared in combined viewpoints in eight teams. The most used combined viewpoint was purpose and activity type.

Table 4-7. Viewpoints that novices used to categorize design techniques

Team No.	Viewpoint	Meaning
3	Activity type	The types that the activities (e.g., observation, simulation) are conducted when using the technique.
15	Phase	The phase that the techniques can be used in the design process.
7, 11	Duration	The time length required when using the techniques
4, 9, 10, 12	Purpose	The reason for using the techniques

1	Phase and activity type	Combine the phases with activity types (e.g., collaborative work, draw chart) when using the techniques
8	Duration and Participant	The time length that different participants need when using the techniques
6	Participant and activity type	The activity types of the involved participants
5	Participant and purpose	The participants' purposes when using a design technique
16, 17, 19	Participant, purpose, and activity type	Combination of participant, purpose and activity type
2, 13, 14, 18	Purpose and activity type	Combination of purpose and activity type
20	Representation of the technique	The superficial appearance of a technique (e.g., bubble chart)

Categories of Design Techniques

When looking at the created categories, the minimum number of the categories was created by team 11 with three categories based on the viewpoint of duration. The maximum number of the categories was created by team 19 with nine categories based on the viewpoints of participant, purpose, and activity type. The average number of created categories was 5.5. As some of the created categories had the same name and some of them had similar meaning, standardized categories were at first created to aggregate the sorting result from novices, which could also help with analyzing created categories and comparing with the experts' created categories. Three researchers went through all categories and created 16 standardized categories from 110 original categories (Table 4-8).

The 16 standardized categories were used by at least two teams. Some categories were built by the single team, for instance, user simulation, user participation (user group), flexible time length, feedback collection (offline), feedback collection (online), user simulation, etc. To analyze the categories created by only one team could also be interesting, but such categories could not present the commonalities between novice. Thus, the standardized categories were mentioned by at least two teams. Jaccard's Coefficient was used to calculate the agreement between teams in the standardized categories because Jaccard's Coefficient is usually used in the analysis of categories from open card sorting (Lewis and Hepburn 2010; Schmettow and Sommer 2016). Compared with the example of standardized open card sorting categories provided in Spencer (2009), the Jaccard's Coefficient showed that the standardized categories could be used for the further analysis of the assignment of design techniques. Around half of the participants created the following categories: user research, idea generation, and information organization. These three categories belonged to the viewpoint purpose of using design techniques.

Table 4-8. Standardized categories and examples of assigned design techniques

Standardized category	#Teams	Jaccard's Coefficient	Standardized category	#Teams	Jaccard's Coefficient
User research	11	0.370	Product evaluation	3	0.400
Idea generation	10	0.340	Collaboration with stakeholders	3	0.405
Information organization	9	0.372	User participation	3	0.489
Feedback collection	5	0.411	Short-term duration	3	0.661
Prototype evaluation	5	0.350	Long-term duration	3	0.474
Prototyping	5	0.400	UX evaluation	2	0.647
Evaluation	4	0.414	Mid-term duration	2	0.528
Expert participation	4	0.551	Relationship and dependency	2	0.536

Note: #Teams means the numbers of teams considered the standardized category.

Assignment of Design Techniques to Categories

The analysis of the standardized categories shows that approximately half of the teams created user research, idea generation, and information organization. Card sorting participants assigned many techniques in the three categories (62, 55, 43 techniques separately). But not all the assigned techniques had a high percentage of the overlap between design techniques and standardized categories. Thus, the percentage of participants that assigned a design technique to a standardized category were computed. In total, there were 70 design techniques and 16 standardized categories. 16 standardized categories and techniques were presented in Appendix E. Take actors mapping as an example, 30% of the 20 teams (i.e., six teams) put the design technique actors mapping in the standardized category information organization. As there were nine teams who created the standardized category information organization (Table 4-8), which meant six out of nine teams allocated actors mapping to standardized category information organization. Thus, actors mapping could be assigned into the category information organization. The entire information was analyzed by following this procedure. When there were more than half of the participants who used the standard categories assigned a technique into the specific standardized categories, the categories were highlighted in Appendix E. As the participants created categories from different viewpoints, the standardized categories were not mutually exclusive. For example, the design technique affinity diagramming was classified into two standardized categories idea generation (eight out of ten teams) and information organization (five out of nine teams).

4.2.4 Comparison

The classification process of novices in the open card sorting presented that the created novice-based classification included viewpoints, categories, and assigned techniques in each category. As the expert- and the novice-based classification needed to be comparable, the expert-based classifications should also include viewpoints, categories, and assigned techniques in each category. Thus, the developed taxonomy in Chapter 4.1 (Liu et al. 2016) and another two classifications (Martin and Hanington 2012; Vermeeren et al. 2010) were used in the comparison. The viewpoints in the novice-based classification can be seen as dimensions in the expert-based classification. The difference is that the dimensions in the expert-based classification are hierarchically structured with the categories, while the viewpoints in the novice-based classification have no parent-child relationship. Moreover, the selected design techniques used for card sorting were also included in these three studies. Thus, it was reasonable to use these three expert-based classifications for the comparison. 11 dimensions from the three expert-based classifications were summarized (Table 4-9).

Table 4-9. Dimensions and categories in existing expert-based classifications

Dimension	Meaning	Categories	References
Design phase	The phase in the design process that the techniques can be applied to	Planning, concepts and early prototypes, functional prototypes, and release and monitor	(Liu et al. 2016; Martin and Hanington 2012; Vermeeren et al. 2010)
Data type	The type of the collected data	Qualitative, quantitative	(Martin and Hanington 2012; Vermeeren et al. 2010)
Evaluation type (i.e. activity type)	The type of evaluation actions when using the techniques	Questionnaire, interview, experiment, observation, group discussion	(Liu et al. 2016)
Duration	Time length when using the techniques	Long-term, short-term, before usage, momentary, single episode	(Liu et al. 2016; Vermeeren et al. 2010)
Participant	People who should participate when using the techniques	User attendance, user absence, specific selection of users, random choice of users, groups of users, UX experts, pairs of users	(Liu et al. 2016; Martin and Hanington 2012; Vermeeren et al. 2010)

Application	The applications that the techniques can be applied to	Web service, mobile service, PC software, hardware design	(Vermeeren et al. 2010)
Origin	The origin of a design technique	Innovative, adapted, traditional	(Martin and Hanington 2012; Vermeeren et al. 2010)
Time dependency	To collect immediate emotional feedback or memory-based feedback	Real-time feedback, retrospective feedback	(Liu et al. 2016)
Content	The content that targeted by the techniques	Behavioral, attitudinal	(Martin and Hanington 2012)
Context	The context that the techniques can be used	Lab, field	(Vermeeren et al. 2010)
Purpose	The purpose of using the techniques	Exploratory, generative, evaluative	(Martin and Hanington 2012)

The comparison process was divided into three steps. *First*, the viewpoints from novices were compared with the experts' created dimensions. *Second*, the standardized categories were used to compare with the categories created by experts. In comparison, whether novices also consider the mentioned categories by experts were presented. *Third*, the result from novices was compared with the mentioned techniques in the experts' categories and presented the similarities and differences in the specific assignment of techniques, which reflected whether the experts-based assignment of design techniques was consistent with the novices' understanding.

Dimensions of Design Techniques

When comparing the collected viewpoints from the card sorting process with the dimensions that are developed by experts, 45.5% of high-level dimensions of design techniques that were developed by experts (Table 4-9) were also considered by novices as viewpoints in the classification process (Table 4-7). The overlapping dimensions between novices and experts included purpose, evaluation type (i.e., activity type), design phase, duration, and participant. Novice-based viewpoints demonstrated that novices tended to use combined dimensions when categorizing techniques. For example, team 16 created six categories: stakeholder, general methods to obtain data, observation, designers enablement, user and designers, and user enablement. The six categories developed were based on three dimensions: participant, purpose, and form. Team 8 created four categories: long-term user participation, short-term user participation, user participation time length can be controlled by designers, and users not participation. The four categories were developed based on the integration of the dimension duration into the dimension participants, whereas dimensions in expert-based taxonomy were independent of each other.

There were dimensions included in the expert-based classifications, but they were not considered by novices. Such dimensions included data type, origin, time dependency, etc. These unmentioned dimensions were not included in the descriptions on the technique cards for the card sorting exercise. Thus, such information cannot be derived by reading the descriptions on the cards but needs the accumulation of design knowledge and experience in design practices. For example, it might be impossible for novices who only attend one or two design-related courses to decide whether a technique is originally invented to design, adapted from other disciplines, or traditionally across discipline. The differences further confirmed that novices tended to classify techniques based on superficial descriptions while experts inclined to classify techniques with the combination of their professional background knowledge and practical experience.

Categories of Design Techniques

When comparing the expert-based categories (Table 4-9) with the novice-based categories (Table 4-8), only 20% of the expert-based categories were included in the novice-based categories. The categories mentioned by experts that related to the novice-based categories include prototyping, release and monitor, long-term, short-term, user participation, and UX experts. These categories were included in the dimensions of phase, participant, duration, and activity type in expert-based classifications. One of the reasons for this result might be that the students that participate in the card sorting exercise already attended some lectures or seminars that introduce design processes. The categorization behavior is tightly related to the existing design knowledge. However, it also demonstrated that the dimensions were not clearly defined in novices' mind. As a result, novices tended to derive categories that were cross-dimensions, which could not be used to explain a specific dimension, whereas the expert-based categories were developed based on a specific dimension, which could be used to explain a specific dimension.

Further differences could be reflected by the categories that were mentioned by novices, but not considered in the existing expert-based categories. For example, more than 25% of participants in the open card sorting procedure developed the categories of information organization and feedback collection, which were not introduced in the existing expert-based classification. When comparing with the existing categories created by experts, each of these two categories had a relatively broad scope. For example, the ways of how information is organized (i.e., information organization) is a representation of a design technique which can be further divided. The category feedback collection is also very generic because as long as a technique collect some information from people, feedback is collected. The novice-based categories were created from a general perspective with a very broad scope, which reflected that novices did not classify techniques as specific as experts.

Assignment of Design Techniques under Categories

As there were differences between novices- and expert-based categories, I also compared the assigned design techniques for each category. In the category idea generation, design techniques aimed to support the creation of new ideas for new products or services or for optimizing existing ones. The design techniques that were included in this category include cognitive mapping (25% agreement), shadowing (25% agreement), etc. (Appendix E). Cognitive mapping and shadowing were categorized in the category concept generation by Martin and Hanington (2012). However, there were also some differences. For example, contextual laddering and co-discovery were categorized in the category idea generation by Vermeeren et al. (2010). The shared novice agreement of classifying these two techniques in idea generation was very low. Novices did not agree to categorize these two techniques in idea generation. In the category prototyping, novices assigned experience prototyping (20% agreement) and collaborative sketching (25 % agreement) (Appendix E), which was consistent with the classification described by Liu et al. (2016). The two keywords "sketching" and "prototyping" in the names of the techniques (experience prototyping and collaborative sketching) could be seen as hints that led to the same assignment of the techniques. The similarities in the assignment of the both mentioned categories by novices and experts presented that when the description of a technique showed a clear relation to the category, novices and experts performed similar categorization result. The comparison also reflected that novices tended to classify techniques based on superficial descriptions.

Looking into the assigned techniques in the categories that were only mentioned by novices showed which techniques were considered specifically. Based on the analysis of categories, information organization and feedback collection were used as examples. The techniques in the category information organization were used for organizing "collect data for further analysis;" for example, cognitive mapping (35% agreement), or actors mapping (30% agreement). One of the reasons could be that a picture was provided to present an example of the usage of the technique on the technique cards. Some of the pictures showed a map of the connected information. Moreover, the two techniques had the same keyword "map." Information organization might be created just because of the descriptions of techniques strongly related to summarizing information. Another example was feedback collection which included photo diary (25%

agreement), UX curve (20% agreement), etc. It was obvious that these two techniques collect data from users. The assignment of design techniques under the categories of information organization and feedback collection confirmed that novices focused on superficial descriptions when classifying design techniques.

4.2.5 Conclusion

In the second part of design cycle 1, an open card sorting procedure with novices was conducted and the collected data were analyzed. Novices who used dimensions, categories, and assigned techniques in specific categories were presented. Subsequently, the result from the open card sorting procedure was compared with existing expert-based classifications. The similarities and differences between novice-based and expert-based classifications were compared. Additionally, the differences in categorization behaviors between novices and experts were also compared.

The introduction of a novice-based classification can be a starting point for the analysis of the effect of different classifications on different people (novices vs. experts) in selecting design techniques. From a practical perspective, the created standardized categories can be used as a novice-based classification. As there was no parent-child relationship between the viewpoints and categories, the novice-based classification includes flat categories without high-level dimensions. One of the forms to present the flat categories can be tags (Conroy et al. 2009; Sinclair and Cardew-Hall 2008).

As the expert-based taxonomy and novice-based tags cannot be directly used for supporting the selection process of design techniques, an artifact is needed to instantiate the created classifications. Subsequently, a further comparison on whether a hierarchically structured classification or a set of flat tags can help novices perform better in selecting design techniques should be conducted. This conclusion leads to the research performed in design cycle 2.

5 Design Cycle 2: Classification-based Decision Support System[1]

Design cycle 2 was based on the results of cycle 1 and was organized in two parts. In part 1, the expert-based taxonomy and novice-based tags were instantiated in a web-based platform. Meta-requirements and design principles were derived as a foundation for the development of the web-based platform. Furthermore, a usability test of the platform was conducted. In part 2, an evaluation of the differences between expert-based and novice-based classifications was conducted in a lab experiment. Figure 5-1 presents a quick overview on cycle 2.

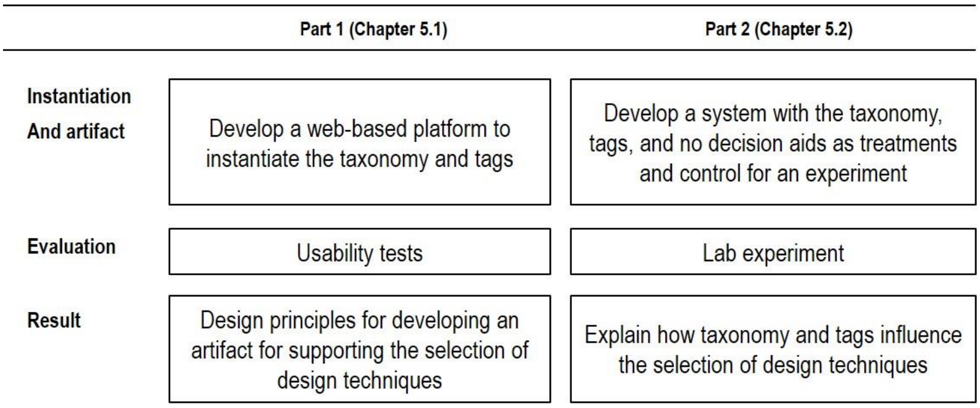


Figure 5-1. Research approach for cycle 2

5.1 A Web-based Platform for Instantiating the Expert-based Taxonomy and Novice-based Tags (ServiceDesignKIT 1.0)

In this Chapter, the web-based platform – ServiceDesignKIT 1.0 – with its associated meta-requirements and design principles is presented. The platform was evaluated in a usability test.

5.1.1 Research Approach

In the first part of design cycle 2, the underlying DSR approach is summarized in Figure 5-2. Based on my understanding of the problem, I derived meta-requirements and design principles. I

[1] This Chapter is based on the following studies which are published or in work: Liu, Leung, et al. (2018); Liu (2018); Liu, Werder, et al. (2018); Liu, Werder, et al. (2019b).

instantiated the principles in the form of a web-based platform. The platform was demonstrated, and a usability evaluation was conducted.

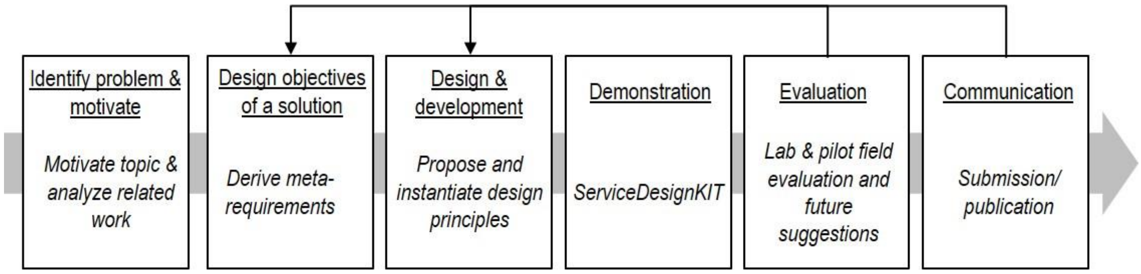


Figure 5-2. Research approach for ServiceDesignKIT 1.0, adapted from Peffers et al. (2007)

5.1.2 Meta-Requirements

The first meta-requirement (MR1) refers to providing a filter for narrowing down the choice of design techniques. As mentioned earlier, the currently used categories in existing web platforms mix the definition of method, technique, and tool, possibly resulting in confusions of terms (Brinkkemper 1996; Kettinger et al. 1997; Sanders et al. 2010). Comparing with the term "design method" and "design tool," a "design technique" provides steps of a procedure and can be classified in several ways for supporting the selection (Brinkkemper 1996; Sanders et al. 2010). Also, with a clear explanation of steps, it is convenient for novices to learn and use design techniques in design processes. Thus, it is necessary for the platform to focus on design techniques. Although it is impossible to include all design techniques, plenty should be included for people to choose from. Moreover, for the purpose of supporting the selection process, it is necessary to include a classification with a limited number of categories to ensure the categories are useful (Miller 1956). Hence, a well-structured classification with useful categories and plenty of design techniques should be included as a basis to support the selection process.

MR1: *Provide a clearly structured classification as a filter to support searching and finding suitable design techniques.*

The second meta-requirement (MR2) refers to enabling co-creation of classifications of design techniques between experts and novices. The key content of the platform is the description of design techniques and their categories. It should be possible to extend and improve over time. As it is impossible for the platform to contain all complete information of each design technique, users should be enabled to edit the existing techniques on the platform. Also, as new design techniques are being developed, it should be possible to add them to the platform. So far, the experts' knowledge-based classifications have not yet been evaluated in the field, and the existing categories of design techniques may not be entirely accurate. Moreover, as more and more novices participate in service design processes, people may have different understandings of categorizing design techniques. A bottom-up suggestion of categories and tags should be enabled because they can reflect people's understanding of design techniques and provide additional access for filtering and searching beyond experts' generated categories (Choi 2015; Golder and Huberman 2006). The co-creation of the classification can provide better support for the selection of design techniques (Nickerson et al. 2013).

MR2: *Enable users of the web-based platform to extend and improve the contained design techniques and bottom-up suggestion of categories and tags describing design techniques.*

The third meta-requirement (MR3) refers to exchanging knowledge of design techniques on the platform. As formal training influences the understanding of design techniques, it is necessary to enable people from different backgrounds to communicate and discuss their understandings of different design techniques (Gerrard and Sosa 2014). Besides, design participants not only need to learn how to use design techniques but also need to collaborate with others (Agid 2016). Moreover, users' comments can further help the optimization of the platform to provide better solutions for filtering design techniques. Hence, the platform should enable people to share their knowledge of design techniques.

MR3: *Enable users of the web-based platform to comment and discuss design techniques.*

The fourth meta-requirement (MR4) refers to personalizing the selection of design techniques. Because the target group is novices, they may have different purposes when using the platform to filter design techniques and different design situations. It should be possible for the users to save their favorite design techniques to a shortlist that will allow them to retrieve and have easier access to these design techniques. A personalizing feature should be included to motivate users' dedication (Kim and Son 2009), which may also motivate them to contribute to the content of the platform.

MR4: *Enable users to save their favorite design techniques to a shortlist.*

5.1.3 Design Principles

In order to address the meta-requirements above, the following design principles (DP) were proposed. Each design principle may solve multiple meta-requirements, and each meta-requirement may be addressed by multiple design principles.

As identified, novices need selection support for finding appropriate design techniques (MR1). Structured categories and plenty of design techniques need to be provided. Because structured categorization can help people's selection process (Bussolon 2009), a taxonomy (i.e., classification) with conceptualized categories needs to be used as a basis for the categorization of the design techniques. The categories of a top-down taxonomy from Liu et al. (2016) (part 1 in cycle 1, Chapter 4.1) and the bottom-up created tags (part 2 in cycle 1, Chapter 4.2) were used on the platform as a filter. In addition, as the platform should include a sufficient range of design techniques for design participants to select from, many design techniques should be provided. So far, 71 design techniques[1] were included. As the initial content may not be complete, users of the web platform should be enabled to add new design techniques, edit the content of existing design techniques and bottom-up suggest categories and tags of design techniques, in order to keep the platform improving and extending (MR2). Ideally, the platform will improve over time in this way. Improved content can also support the selection by offering better content that helps users distinguish and decide among the provided design techniques. A control system should be included in order to allow admins to check and approve changes to the content (Chaturvedi et al. 2011). This feature is intertwined with the bottom-up suggestion to make sure the suggested contents are checked before publishing to the public.

DP1: *Include a pre-defined top-down structured taxonomy as well as bottom-up suggested tags of design techniques and a controlled knowledge base for users to suggest new design techniques and the categories to improve the web-based platform.*

A communication capability is needed for fostering discussion (MR3). Professional designers can share their experiences with specific design techniques, which can help others to find suitable design techniques. Novices can ask questions by leaving comments. The knowledge exchange is consistent with the emphasis on collaboration between experts and novices (Hileman 1998). This feature can further enhance the content of the platform (MR2). In order to motivate people to use

[1] The initial version had 70 design techniques from the literature and websites, which provide a detail description of each. One of the users suggested a design technique, which made it become 71. The number will change in the future.

the platform, personalization features should be considered. Users can access a shortlist of favorite design techniques which are suitable for their purposes and design situations (MR4). A benefit of this feature is that recurrent users can save their favorite design techniques and retrieve them on their next visit. Another benefit is that, after saving a shortlist, users can filter appropriate ones from a small group of potentially suitable ones and further narrow down the selection result.

DP2: *Include personalization features for users to save their favorite design techniques and communicate with others.*

A short summary of the two proposed design principles and associated meta-requirements is presented in Table 5-1.

Table 5-1. Design principles and associated meta-requirements for ServiceDesignKIT 1.0

DP	MR addressed by DP
DP1: Provide classifications and controlled users' suggestion of design techniques	MR1, MR2
DP2: Communication and personalization features	MR2, MR3, MR4

5.1.4 Prototypical Implementation

On the basis of the proposed design principles, a web-based platform was developed – ServiceDesignKIT 1.0. Figure 5-3 presents a screenshot.

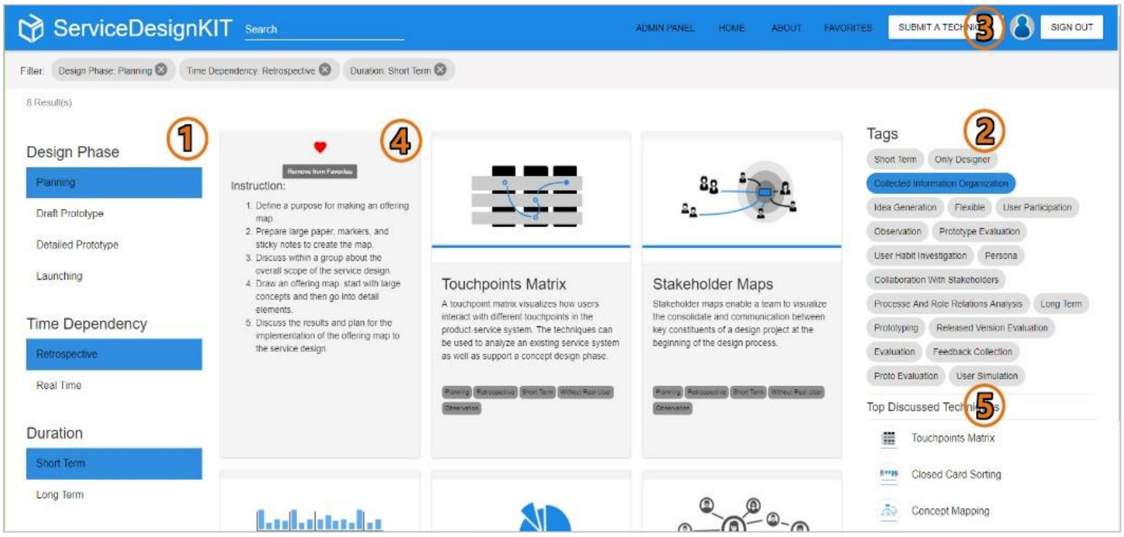


Figure 5-3. Homepage of ServiceDesignKIT 1.0 (Numbers refer to specific features)

The filter function was realized by using taxonomy and tag cloud (DP1). On the left-hand side, an experts' knowledge-based taxonomy was provided (1 in Figure 5-3); on the right-hand side, a bottom-up built tag cloud was included (2 in Figure 5-3). Users were not limited to using only one filter but could mix the two. In order to provide plenty of design techniques for users to select, suggestions of new design techniques were enabled. When clicking on the button for submitting a design technique (3 in Figure 5-3), a new design technique could be added, which enabled the improvement of the content of the web platform (DP2). The appearance of each design technique was designed as a flip card (4 in Figure 5-3). The use of the flip cards, in which digital objects behaved like their counterparts in reality, was shown by the easy transition from front to back. By hovering over a flip card, people could switch from the front to the back. On the back of the flip card, there was a heart symbol which enabled people to add the design technique to a shortlist for retrieving (DP2). When clicking on a design technique card, a subpage with detail information of each design technique would appear. Users were enabled to communicate with others by leaving comments, giving feedback and asking questions (DP2). The design techniques that were most discussed were depicted below the tag cloud (5 in Figure 5-3). In the subpage of each design technique, users could edit the content, suggest categories and tags of existing design technique. On submission, an edited design technique ran through the content management system that was only accessible to the platform’s administrator. Pending design techniques were temporarily inaccessible for users of the platform. The administrator could check, approve, edit, and remove the newly added or edite design techniques to ensure the quality of the content (DP1).

5.1.5 Evaluation

The evaluation of ServiceDesignKIT 1.0 was separated into two stages. The first stage was a usability test using eye-tracking and retrospective think-aloud. The second stage was a small-scale pilot study with the application of ServiceDesignKIT 1.0 in a Master course for KIT students.

Usability Test

The purpose of the usability study was to know how people used ServiceDesignKIT 1.0 before they became familiar with it. Eye-tracking and retrospective think-aloud are relevant techniques for performing usability tests. Eye-tracking can be used to collect people’s first impression and cognitive process of websites (Duchowski 2002; Nielsen and Pernice 2009). Retrospective think-aloud enables experiment participants to work in silence and describe their experience and thoughts afterward (Fonteyn et al. 1993; van den Haak et al. 2006). Thus, eye-tracking and retrospective think-aloud were combined to evaluate whether the features on ServiceDesignKIT 1.0 could be recognized by people.

Six people participated in the usability test. Because three to four participants can yield 80% of the usability findings (Lewis 1994; Virzi 1992), six participants were sufficient for this usability test. The participants included four males and two females whose ages ranged from 25 to 34. All of them were pursuing or already had a higher university education (four Ph.D. students and two Master students, with backgrounds relevant to information systems, software engineering). None of them was familiar with ServiceDesignKIT 1.0. Three tasks were prepared for participants: i) to select up to three design techniques and add them to the favorite list; ii) to find the top-discussed design techniques and write a comment; iii) to submit a new design technique with the given attributes. A detailed description of the tasks is in Appendix F. Before giving the tasks, participants were asked to imagine that they had to create a new service and planned to select design techniques on the web-based platform. During the experiment, the participants at first looked at each web page for five seconds and got the first impression. The eye-movement data were recorded by a commercial web-based eye-tracking application (eyezag.de) (Zugal and Pinggera 2014). Then, the participants were asked to use ServiceDesignKIT 1.0 to perform the tasks. The screen was recorded during the process. After they finished, they were asked to verbalize their thoughts while watching screen recordings. Their descriptions were recorded for analysis.

The evaluation result showed that all participants could understand the main features of ServiceDesignKIT 1.0. They agreed that the categorization of the design techniques was clear, and the function of editing design techniques, adding to the favorites, and giving comments were working well. Overall, ServiceDesignKIT 1.0 seemed to instantiate the proposed design principles successfully. However, there were issues identified, which may need to be considered in further optimization.

For a detailed analysis of visual activities of participants during the eye-tracking study, different areas of interests (AOIs) on the homepage were defined (left of Figure 5-4). AOIs are defined as areas of a display or visual environment that are of interest to the research (Hyrskykari et al. 2008). During the eye-tracking study, fixations of the participants were recorded. Figure 5-4 (right) presents the heat map of the homepage. In a heat map, cold colors such as blue indicate fewer fixations, while the warm colors like green, yellow, and red indicate more fixations. Based on collected fixations and the provided heat map, the illustrations of design techniques attracted the attention of the users more than the other parts. In Table 5-2, the detailed results of the eye-movement data were summarized with the information of hit rate (ratio of participants that were fixed inside the AOI), average fixation duration, and the average time to the first fixation. The recorded eye-movement data showed that design technique cards (AOI2) attracted the most attention (all of the participants with an average duration of 2.81s) in comparing with the other AOIs. Categories of design techniques (AOI1) also drew a lot of attention (83.33% of the participants with an average duration of 0.84s), while the top discussed techniques (AOI4) got the least attention. The time to the first fixation presented that the first noticed AOI was designed technique cards (AOI2), and the last noticed AOI was the top discussed techniques (AOI4). The submission function (AOI5) was an exception since the location was far from the other four AOIs. Only one participant recognized AOI5 (with a fixation duration of 0.77s and time to the first fixation of 0.33s).

Notes: AOI1 is the filter function; AOI2 includes design techniques that are presented on flip cards; AOI3 displays the tags; AOI4 includes top discussed design techniques; AOI5 includes functions for submitting a new design technique and viewing favorites.

Figure 5-4. Left: AOIs of the homepage. Right: heat map of the homepage.

Table 5-2. Statistics of AOIs

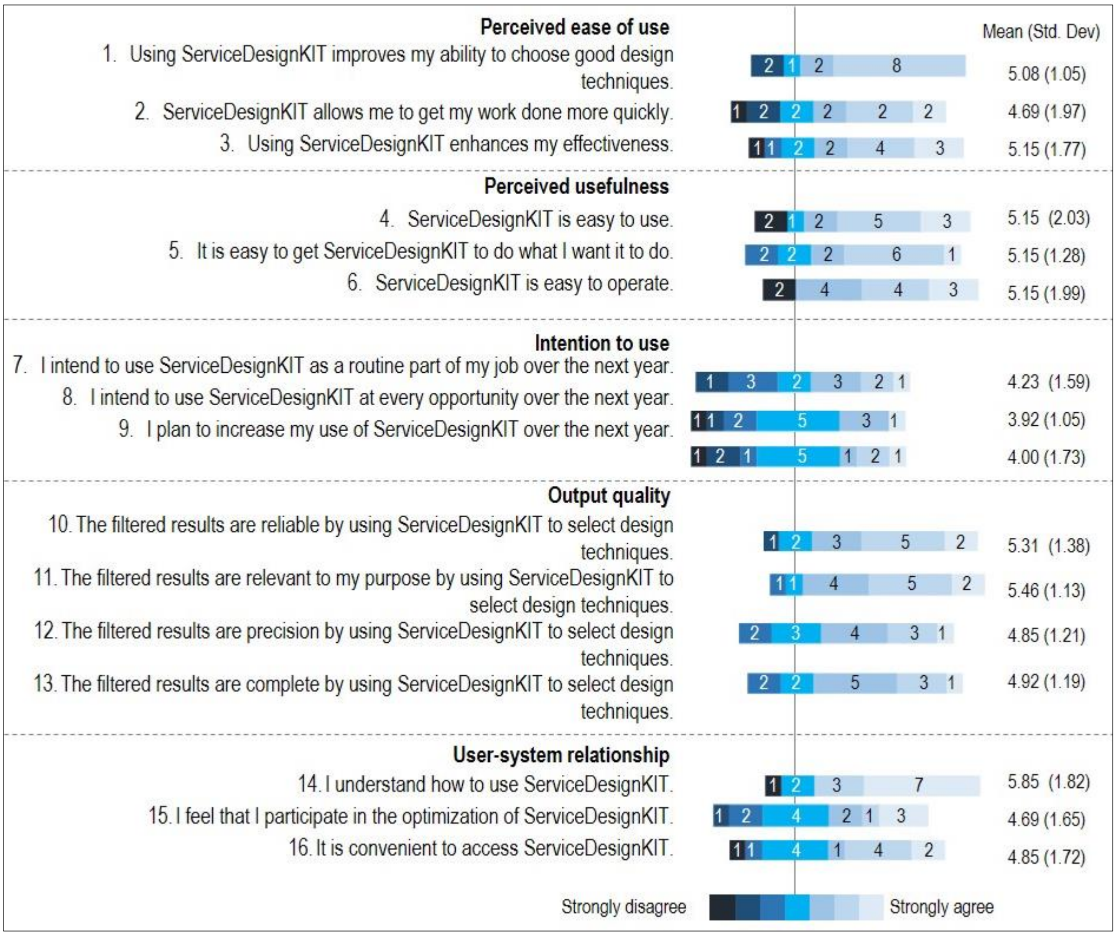
	AOI1	AOI2	AOI3	AOI4	AOI5
Hit rate	83.33%	100%	33.33%	16.67%	16.67%
Average fixation duration	0.78s (SD=0.55s)	2.81s (SD=0.64s)	0.84s (SD=0.44s)	0.13s (SD=0s)	0.77s (SD=0s)
Average time to first fixation	1.39 s (SD=1.28s)	0.02 s (SD=0.02s)	3.44 s (SD=0.97s)	4.44 s (SD=0s)	0.33 s (SD=0s)

The combination of the results from eye-tracking with the results from retrospective think-aloud demonstrated that people paid attention to the design features implemented design principles when using the web-based platform for their tasks. For DP1, participants could understand the categories as a filter, as participant3 said: "*That's what I discovered, I could actually draw down and use more than one category*." The feature of bottom-up suggestion of design techniques was also obvious, e.g., "*The next was to submit a technique. It was very nice that I clicked on that directly (Participant4).*" Additionally, the participants were satisfied with the visualization design. For example, flip cards components of design techniques (AOI2), participant1 said: *"I* *think they are pretty nice the cards getting flip."* For DP2, the comment feature was straightforward to use, as participant5 said "*I clicked on one of the top discussed techniques and I saw the comment box. I added my comment. This step is very easy.*" The personalization feature was also well addressed by providing a favorite list, e.g., "*I clicked on the heart and added it to my favorites (Participant2)."* However, some detail features still could be optimized. First, the left-side categories could be visualized more like a filter. Although the categories drew participants' attention (AOI1) and could be used to filter design techniques, they were not as self-explaining as expected, as participant3 told us "*But I didn't know that they are something like a filter.*" Second, the top discussed techniques were not easy to be recognized. The eye-tracking result presented top discussed techniques (AOI4) attracted little attention, which was coherent with the participants' feedback in retrospective think-aloud section. Participant3 said, *"it was quite difficult to find top discussed techniques because it wasn't highlighted much."* Third, the submission function (AOI5) was not very obvious, but people could find it when they needed it. The optimization suggestion of the submission function related to showing the feedback after submitting a design technique.

Pilot Study

A first attempt was to use ServiceDesignKIT 1.0 in the winter semester of 2017 within the lecture "Digital Service Design" offered at KIT. ServiceDesignKIT 1.0 was provided as a supplement for students to select and apply design techniques as part of the exercises of an applied capstone project. ServiceDesignKIT 1.0 was introduced in the lecture, but students were not forced to use it to select relevant design techniques. 33 Master students (three to four build up a team, nine teams in total) participated in the capstone project. The capstone project was about a design challenge in the field of financial services given by one of our industry partners. It focused on creating a prototype of a Robo-Advisor. Each student team used several design techniques to address the challenge and delivered a low-fidelity prototype. In the last lecture, a questionnaire was distributed to the students who attended the lecture, and 13 complete questionnaires were collected. The average age of the students who filled out the survey was 25 years, and five of them were female, eight were male. They studied industrial engineering, as well as information engineering and management. The students could be seen as novices in digital service design, who might be the potential users of ServiceDesignKIT 1.0 in the future.

The questionnaire for the pilot evaluation was set up based on the design principles. Output quality, user-system relationship, perceived ease of use, perceived usefulness, and intention to use were measured. User-system relationship refers to a user's knowledge of the system, involvement, and access to the system (Ives et al. 1983; Raymond 1985). Output quality refers to the relevance, accuracy, precision, and completeness of the output from the system (Ives et al. 1983; Mirani and King 1994). Intention to use relates to the extent that people intend to use a system in the future as a routine part of the job at every opportunity (Wixom and Todd 2005). Perceived ease of use and perceived usefulness refer to user-friendly and the enhancement of job performance (Wixom and Todd 2005). Each question was measured on a 7-point Likert type scale (1= strongly disagree, 7 = strongly agree). Excel 2016 was used to analyze the mean and standard deviation of each item (Cronbach's Alpha = 0.875). The detailed result is presented in Figure 5-5.



Note: The number in the stacked bar chart means the number of participants that chose a level of agreement.

Figure 5-5. The result of field evaluation.

The evaluation of perceived ease of use and perceived usefulness showed that the participants agreed that the web platform could enhance the effectiveness of selecting design techniques and was easy to operate. Figure 5-5 shows that most of the participants had positive perceptions with regards to ease of use and usefulness. The output quality also received a high evaluation, which meant the filter results by using the web platform could be considered to be trustworthy by the users. Furthermore, it explained that the implemented top-down classification in the platform could provide reliable, relevant and precise filter results and the bottom-up suggested tags could also support the selection process (DP1). When looking at the means of the questionnaire, the evaluation result of the user-system relationship was quite similar to output quality. However, the result presented that there were fewer participants who held a positive attitude on the user-system relationship than output quality. Although the result reflected the easy access of the web platform, much effort needed to be put into motivating people to participate in the optimization of the platform (DP2). Comparing to the mentioned four constructs, the means of items of intention to use were relatively quite low. The stacked bar chart presented that the ratio of participants who gave positive and negative evaluations to intention to use was nearly the same. One reason could be that the web platform needed more features to motivate people to use it in the long run. Another reason might regard to the sample. Some students might only use this tool for the lecture but not have a further career plan on the domain of digital service.

5.1.6 Conclusion

The first part of design cycle 2 proposed an artifact based on the classifications (i.e., expert-based taxonomy and novice-based tags) of design techniques created in cycle 1. The derived meta-requirements and suggested design principles provided guidance for researchers to instantiate web-based platforms supporting the selection process of design techniques. The suggested design principles sought to use the top-down classification from the experts' perspective and the bottom-up suggested tags from novices' perspective as a filter. This attempt was a starting point to analyze differences and similarities of top-down and bottom-up classifications. In addition, the attempt of applying different classifications was a basis for further research on the effects of classifications on novices' selection of design techniques. Practically, ServiceDesignKIT 1.0 could be used as a filter for the selection of design techniques under different design situations. Besides a top-down classification, ServiceDesignKIT 1.0 enabled people to suggest bottom-up tags of design techniques, which could reflect novices' understandings of design techniques and emphasize the importance of the involvement of novices in digital service design processes. ServiceDesignKIT 1.0 enabled people to create and improve information about design techniques and also prevented the contribution of inaccurate information. The use of controlled knowledge base and the implementation of communication capabilities in the form of user comments enabled the refinement of the categories and extension of design techniques included in the platform, which could further improve the content of ServiceDesignKIT 1.0.

5.2 The Effect of Taxonomy and Tags on the Selection of Design Techniques

The expert-based taxonomy and novice-based tags in the web-based platform are two types of decision aids which are interventions that should help in the decision-making process (Elwyn et al. 2010). As part of the usability test of the web-based platform, the performance of these two decision aids were not evaluated. Thus, in this Chapter, a systematic analysis of the influence of taxonomy and tags on the novices' selection of design techniques is conducted in a lab experiment. The overall research approach for this experimental study is described in the following. Hypotheses and the experiment design are introduced in detail. Finally, the collected data is analyzed and discussed.

5.2.1 Research Approach

The overall research approach is presented in Figure 5-6. Before conducting the experiment, a set of hypotheses was proposed. The experiment tool was developed with embedded taxonomy, tags, and control. The experiment was conducted with 195[1] participants (invited from KIT Hroot[2]) in 12 sessions in three days in the KD2lab. The data from the experiment was analyzed in order to test the hypotheses. Subsequently, the results were presented, discussed, and reflected.

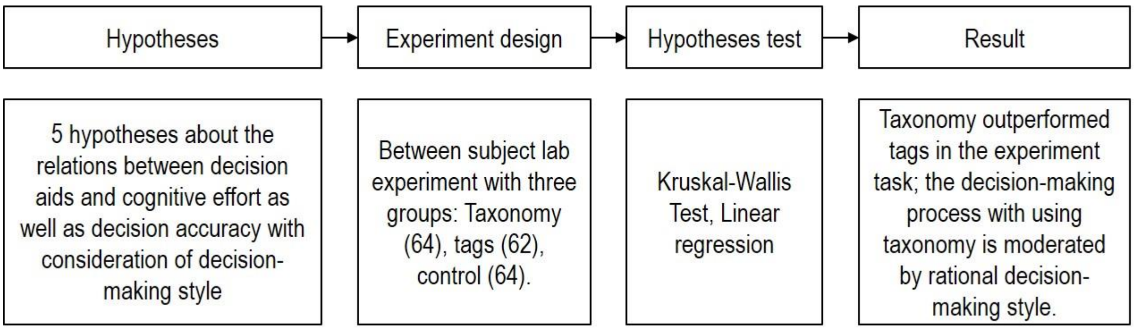


Figure 5-6. Overall research approach for comparing taxonomy and tags

5.2.2 Hypotheses

Based on cognitive fit theory, taxonomy and tags can be seen as two different problem-solving representations. The task is selecting design techniques for specific design scenarios. The task

[1] 190 participants provided valid data for the experiment analysis. Five data points were removed because of incomplete or misunderstanding of the experimental tasks.

[2] https://iism-kd2-hroot.iism.kit.edu/?locale=en

performance of using decision aids can be measured by the selection accuracy. In the decision-making process, system cognitive effort is generated while using decision aids. Based on cognitive fit theory, it can be assumed that when the decision aids match the selection task, the system cognitive effort will decrease, and the selection accuracy will increase. Additionally, cognitive theory also introduces that a consistency between mental representation and the problem-solving process can lead to a better performance because there is no need to transform the mental representation to adapt to the problem-solving representation (Vessey 1991; Vessey and Galletta 1991). Thus, the decision-making style is assumed to have a moderation effect between the relationship of decision aids and selection accuracy. In the following, the individual hypothesis is introduced.

The Effect of Decision Aids on System Cognitive Effort

When performing a task, system cognitive efforts are generated by using a system (Pereira 2000). In a decision-making task, scanning the information of all alternatives and comparing the similarities as well as differences causes system cognitive efforts (Bettman et al. 1990). Thus, the more selection alternatives exist, the higher system cognitive effort will be generated. One of the ways to reduce system cognitive effort is to present less information to decision-makers (Iyengar and Lepper 2000). This can be realized by using decision aids to narrow down the selection scope and reduce the selection alternatives (Tremblay et al. 2010). Hence, the use of decision aids should result in a lower system cognitive effort.

When making decisions with the help of a decision aid, the match between the decision-makers' preferences and the selection conditions provided in the decision aid plays a key role in processing information, because of a mismatch in between leads to confusion which can cause system cognitive effort (Engin and Vetschera 2017; Hong et al. 2004). When novices select appropriate design techniques, extra system cognitive effort can be created by unclearly defined categories. Compared with tags, categories in a taxonomy are hierarchically organized, this should support understanding the categories to select the right information. Tags are expected to create higher system cognitive effort than a taxonomy. Thus, using a taxonomy leads to lower system cognitive effort than using tags. Thus, the following hypothesis is proposed:

> ***H1:*** *The decision aid taxonomy will lead to lower system cognitive effort than decision aid tags*

The Effect of Decision Aids on Selection Accuracy

As too-much-choice increases choice complexity, providing a large amount of alternatives leads to difficulties when making decisions (Bollen et al. 2010; Greifeneder et al. 2010) and negatively influences selection accuracy (Johnson and Payne 1985). On the contrary, limited choices can lead to greater satisfaction (Iyengar and Lepper 2000). As the number of design techniques is increasing, more selection choices will appear. However, too-much-choice effects can be reduced by narrowing down the selection scope (Greifeneder et al. 2010). In order to reduce the selection scope, decision aids can be used to assist the decision process by filtering the selection possibilities based on certain decision conditions (Tremblay et al. 2010). The selection of design techniques can be based on the match of novice designers' design conditions and the classified categories of the design techniques. Following cognitive fit theory, when the problem representation matches the task, the performance should improve (Vessey and Galletta 1991). In this case, the decision aids can be seen as the problem representation; the selection accuracy is an expression of the performance. Because the decision aids match the task of selecting design techniques, the selection accuracy should increase.

As a form of decision aids, a classification can be created by experts or novices. As experts have rich background knowledge, the categories are hierarchically structured, whereas novices make categorizations more randomly without fully thinking about the parent-child relations between the categories (Schenk et al. 1998). Although both types of classifications can positively contribute to the selection process by narrowing down the selection scope, the expert-based taxonomy is more precise than the novice-based tags. Existing research in knowledge organization shows that a clear knowledge structure can help with the task performance (McKeithen et al. 1981). Although using tags may better suit novices' understanding than using a taxonomy, the selection task itself needs a structured understanding of the content (Chi et al. 1981). For example, existing research on multi-criteria decision aid reflects that the decision aid needs to be logical, clear, and well-structured in order to support decision-makers when comparing and evaluating the options (Thawesaengskulthai and Tannock 2008). However, generating properly structured categories is difficult for novices (Chi et al. 1981). Decision aids are suggested to help novices structure information in the decision-making process (Mackay and Elam 1992). Providing a high structured classification for novices in a selection task can complement the shortage of lacking this skill. Hence, the decision aid taxonomy is assumed to outperform tags in selection accuracy. Thus, the following hypothesis is proposed:

H2: *The decision aid taxonomy will lead to a higher selection accuracy than decision aid tags*

The Relation between System Cognitive Effort and Selection Accuracy

In any decision-making tasks, system cognitive effort is produced inevitably when processing information (Pereira 2000). This has a negative effect on the task performance (Garbarino and Edell 1997). For example, system cognitive efforts can distract a person's attention from the task to understand how to use the tool (Johnson 2010). The distraction is a noise factor to the main task (Robinson et al. 1997). The efforts cost for thinking, understanding, and figuring out how to use the decision aids can cause a reduced focus on the decision-making task (Johnson 2010). Reducing the system cognitive effort during the task performance has a positive effect on selection accuracy (Einhorn and Hogarth 1975; Todd and Benbasat 1999). Thus, when novices selecting design techniques for a specific design task, the less system cognitive effort created with the decision aids, the more attention can be paid to the selection of appropriate design techniques. Consequently, the focus is mainly set on the decision-making task and more attention can be concentrated on comparing and understanding design techniques. Finally, the selection accuracy will be increased. Thus, the following hypothesis is articulated:

H3: *A reduction of system cognitive effort will lead to an increase in selection accuracy*

Moderating Effect of Decision-making Styles

Decision-making style refers to "the typical manner by which individuals make decisions" (Hamilton et al. 2016) which is generated based on cognitive thinking styles and can be roughly distinguished as rational and intuitive (Allinson and Hayes 1996). In a decision-making task, people with a rational decision-making style are analytic and tend to structure information based on evidential factors such as systematically gathering information as well as considering the alternatives (Epstein et al. 1996; Hamilton et al. 2016). Following different approaches (i.e., holistic or analytic) in the decision-making process can influence task performance (Armstrong 2000; Engin and Vetschera 2017). The rational decision-making style can influence the process of processing structured information. For example, people with a rational decision-making style tend to understand and search structured information when using self-service technologies (Simon and Usunier 2007). Taxonomy suits the analytical style as it includes the hierarchically organized categories (Nickerson et al. 2013). Hence, the influence of decision aid taxonomy toward accuracy is expected to be stronger for individuals with rational decision-making style. Thus, the following hypothesis is articulated:

> ***H4a:*** *Rational decision-making style has a positive moderating effect on the influence between decision aid taxonomy and selection accuracy*

In contrast to the rational decision-making style, the intuitive decision-making style drives an unstructured and randomized information processing process in a decision-making task (Epstein et al. 1996; Hamilton et al. 2016). Intuitive decision-makers are holistic and tend to follow their feelings and hunch, which is different from using structured information (Epstein et al. 1996; Hamilton et al. 2016). However, the unstructured decision-making process is similar to the creation process of tags. Tags suit the holistic style as the tags are at the same level and created based on the surface information (Goh et al. 2009; Golder and Huberman 2006; Klašnja-Milićević et al. 2017). Hence, the influence of the decision aid tags toward accuracy is assumed to be stronger for individuals with intuitive decision-making style. Thus, the following hypothesis is articulated:

> ***H4b:*** *Intuitive decision-making style has a positive moderating effect on the influence between decision aid tags and selection accuracy*

Based on the hypotheses developed above and framed into the input-process-output framework (Bhattacherjee 2012), the research model is depicted in Figure 5-7. Decision aids are inputs for novices' selecting design techniques. The output is the selection accuracy. The process includes the generation of system cognitive effort and the consideration of rational as well as intuitive decision-making styles.

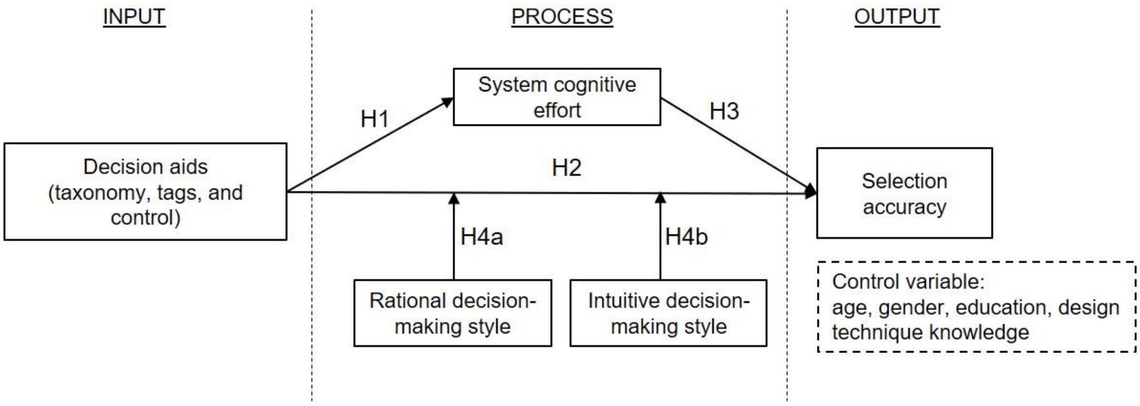


Figure 5-7. Research model for the hypotheses of the effect of taxonomy and tags on selection

5.2.3 Lab Experiment

In order to test the hypotheses and the research model, a between-subject lab experiment was conducted, two treatments (taxonomy vs. tags) and a control (no decision aid) were used. In the following, experiment design and experiment participants are introduced.

Experiment Design

Experiment Setting

An experimental tool that instantiates the treatments and the control group was specifically developed for the experiment. For the treatment of the decision aid taxonomy, an existing taxonomy of design techniques was used (Figure 5-8) (part 1 in cycle 1, Chapter 4.1) (Liu et al. 2016). The taxonomy consisted of high- and low-level categories which provided a clear structure of all the categories. By choosing a relevant category, the search scope was narrowed. When more relevant categories were chosen, the selection scope could be reduced. The treatment of a decision aid tags was created by 40 MSc. students using card sorting (part 2 in cycle 1, Chapter 4.2) (Figure 5-9). Unlike the taxonomy, tags had no parent-child relations, but the design techniques were also categorized, and the categories could also be used to reduce the selection scope. The control group just received a list of design techniques without any decision aid (Figure 5-10). Participants had to click on the corresponding technique and read the provided information for deciding which one to choose. The included content of design techniques[1] was the same between the treatments and the control.

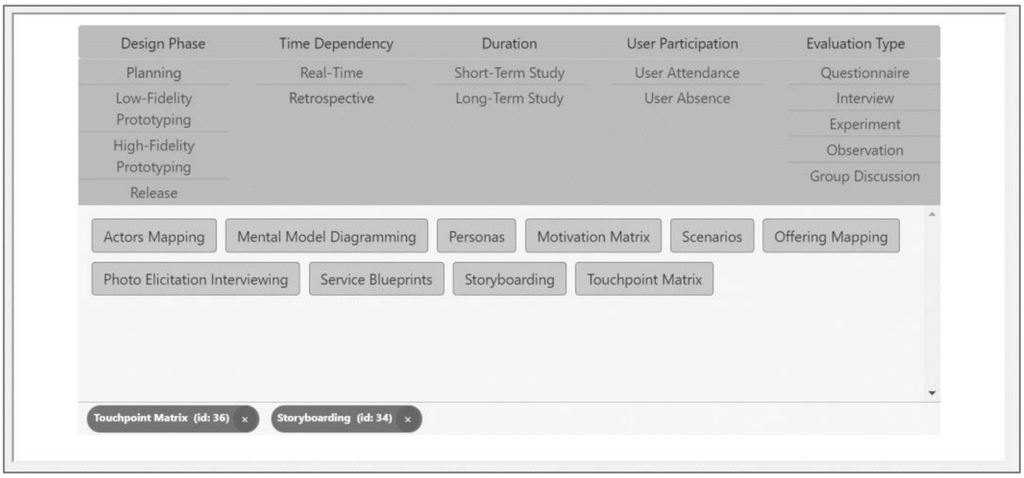


Figure 5-8. A taxonomy of design techniques (treatment of decision aid taxonomy)

[1] The taxonomy, tags, and techniques used in the experiment can be found in Appendix D, E, and G.

Figure 5-9. Tags of design techniques (treatment of decision aid tags)

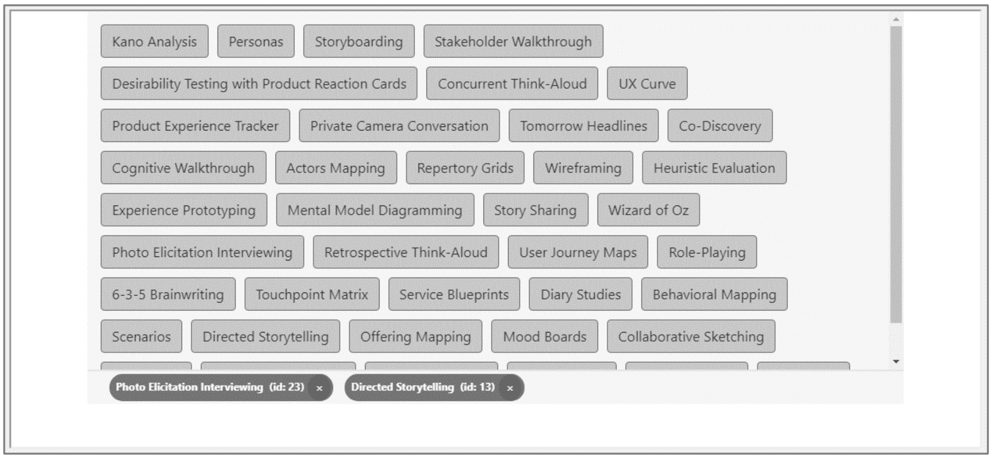


Figure 5-10. A list of design techniques (control)

The tasks used in the experiment were contextualized in a mobile app design project. Three scenarios which refer to three design stages (generating ideas, creating prototypes, and evaluating app) were provided. The tasks were the same between the treatments and the control. Each participant should select maximum five most appropriate design techniques based on the understandings for each of the three scenarios. The selected design techniques were recorded for analyzing accuracy. For each scenario, there were seven appropriate design techniques as correct selections. The description of the three task scenarios and the corresponding design techniques from the literature can be found in Appendix G. In order to make sure the participants understood what a taxonomy or a set of tags is and how to use the tool to perform the task, a short training before the experiment task was provided. In the training stage, examples of classifying animals and finding animals by using a taxonomy or a set of tags were used in order not to give any information to experiment participants on design techniques or the task scenarios. Using such generic examples in the training stage can avoid bias from the preconceived information on the experiment tasks before presenting the task scenarios to the experiment participants.

Experiment Procedure

In order to make sure each participant can conduct the experiment without any interruptions from the experimenters or the other participants, the experiment should be as automatic as possible. The tool Limesurvey[1] was used. The experiment tool with the treatments and control were embedded into Limesurvey. The whole experiment procedure is presented in Figure 5-11. The procedure was separated into three stages after the overall introduction with a description of the data privacy and the experiment stages to the participants.

In the *first* stage, the definition of the concept of a design technique was introduced. As most of the experiment participants were novices, an example of a design technique was used instead of direct introducing a definition. Besides the example, under what situations the technique could be used was also introduced. The description and example of the design technique had no relation to the experiment task scenarios in the second stage. After reading the information, the participants required to answer whether they understood the concept of design techniques. This question was also used as a foundation to make decisions for removing data points for the analysis. After introducing the concept of design techniques, a taxonomy or tags were described to participants by using videos and examples instead of abstract definitions. Sorting examples from zoology were provided. In the first video (around 1 minute), several animal pictures were presented to the participants and classified into a taxonomy or tags. Subsequently, the second video (around 1.5 minutes) showed how to use the created taxonomy or tags to find animals based on certain filter conditions. The screenshot of the videos used in stage one can be found in Appendix H. Following each video, participants had to answer whether they understood the description. If a participant did not understand the information, the data point had to be dropped.

The *second* stage was the main experiment task. Three task scenarios were included. After introducing the whole task, each scenario task was presented on an individual webpage with the embedded experiment tool (a taxonomy, tags, or a list of design techniques). Between two task scenarios, there was a one-minute washout phase to make sure the participants had a short break and the answer to the next task scenario wouldn't be influenced by the previous one. A nature picture was presented in the washout phase. The complete description of the task scenarios is in

[1] "LimeSurvey - The No.1 of Open Source Survey Tools" https://www.limesurvey.org/

Appendix G. Participants' selected design techniques for each scenario were recorded for the analysis of the accuracy.

The *third* stage of the experiment included several questions on self-evaluation of the experiment task. Measurements such as system cognitive effort (Hong et al. 2004) and task motivation (Erez and Judge 2001) were used. In the personal question, the measurement of decision-making styles (Hamilton et al. 2016) was used. Demographic questions were also asked in this stage.

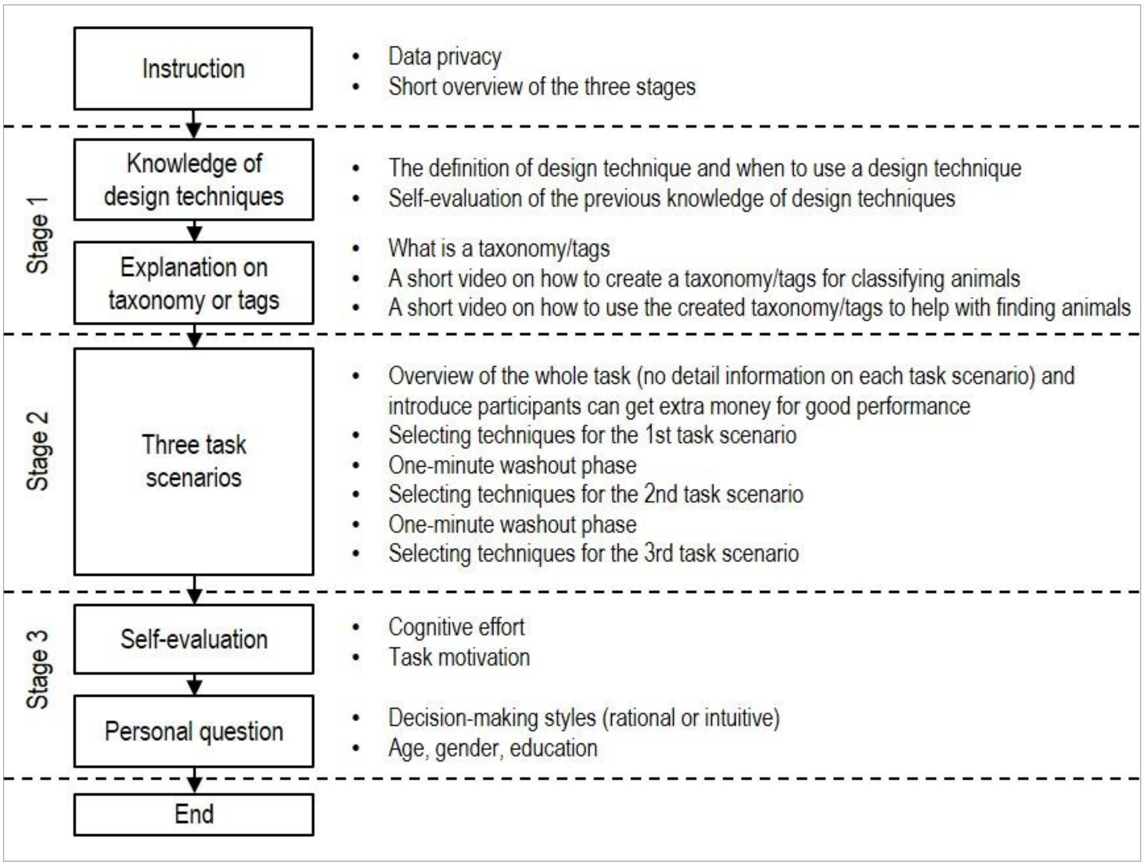


Figure 5-11. Experiment procedure for the evaluation of taxonomy and tags

Operationalization

As the hypotheses focus on design novices, perceived design technique knowledge was used as a control variable (Alexander et al. 1991; Brucks 1985; Pereira 2000). The treatments were decision aid taxonomy and tags. The rational and intuitive decision-making styles (Hamilton et al. 2016) are independent variables which were measured by using a 7-point Likert scale (Burton-Jones and Straub 2006) (Appendix I). This measurement introduced in Hamilton et al (2016) was chosen because the measurement specifically focuses on decision-making styles instead of a generic cognitive style (e.g., Allinson and Hayes 1996; Martin 1998). Moreover, this measurement has been used in decision-making research (e.g., Pajala 2019; Qahri-Saremi et al. 2018). As the experiment focused on the support of decision aids, the measurement introduced in Hamilton et al. (2016) was considered as appropriate. The dependent variables comprised cognitive effort and selection accuracy. The system cognitive effort was measured based on Pereira (2000) by using a 7-point Likert scale (Burton-Jones and Straub 2006) (Appendix I). This measurement focuses on the decision-making task and has also been using in decision-making research (e.g., Hong et al. 2004; Wang and Benbasat 2009). Unlike system cognitive effort, the dependent variable selection accuracy was measured by checking the accuracy of each participant selected design techniques for three task scenarios. Based on literature review, for each task scenario, there were seven correct selections. The number of correct selections were counted as the score of the selection accuracy. As there were three task scenarios, and maximum of five techniques were allowed to be selected for each scenario, the score of the accuracy ranged from 0 to 15. A summary of the decision aids as well as variables and definitions are presented in Table 5-3 and 5-4.

Table 5-3. Definitions of two decision aids

Decision aids	Definition	Source
Decision aid taxonomy	A taxonomy is created by experts which are hierarchical structured and consists of high-level dimensions and low-level characteristics	(Nickerson et al. 2013)
Decision aid tags	A set of tags is created by a group of users of the information without a hierarchical structure.	(Golder and Huberman 2006; Spiteri 2007)

Table 5-4. Definitions of each variable

Variable	Definition	Source
System cognitive effort	System cognitive effort is the mental costs on understanding the system for performing a task.	(Pereira 2000)
Rational decision-making style	The rational decision-making process is based on a systematic evaluation of all alternatives.	(Hamilton et al. 2016; Phillips et al. 1984; Scott and Bruce 1995)
Intuitive decision-making style	The intuitive decision-making process is based on hunches and feelings.	(Hamilton et al. 2016; Phillips et al. 1984; Scott and Bruce 1995)
Perceived design technique knowledge	How much people think he or she has the knowledge about design techniques;	(Alexander et al. 1991; Brucks 1985; Pereira 2000)
Selection accuracy	The number of selected design techniques that are appropriate for the specific task scenario.	-

Participants and Incentives

195[1] people (invited from KIT Hroot[2]) were recruited to participate in the experiment to make sure there were enough subjects in the experiment. After conducting the experiment, five data points were removed because of incomplete or misunderstanding of the experimental tasks. Thus, 190 participants provided valid data for the experiment analysis (64 for taxonomy, 62 for tags, 64 for list). 56 were female, and 134 were male. The average age was 23 years (SD = 4.2). 97% of the participants were students. It took on average 44.8 minutes (SD = 9.8) for the whole experiment (i.e., including all three stages). As the purpose of this study was to understand whether novices could benefit from decision aids, experienced design techniques experts were not necessary to participate in the experiment. Students were seen as an appropriate group for the experiment (Druckman and Kam 2012). Each participant was assigned to one of the three treatments randomly. Each participant conducted the experiment tasks in individual soundproofed computer cubicle in a well-facilitated lab. 12 experiment sessions were conducted in three days. Each participant received at least 8 Euros for the participation. As there were correct answers in the experiment tasks, the participants could get a maximum of 3 Euros extra for their good performance. Experiment participants were given additional incentives to motivate them to perform to the best they could in the experiment tasks.

5.2.4 Data Analysis and Results

Measurement Validity

An exploratory factor analysis using principal component analysis and varimax rotation was performed on system cognitive effort, rational decision-making style, intuitive decision-making style, and perceived design technique knowledge (cf. Yi et al. 2015). Results showed that items loaded highly on their intended factor and lowly on the other factors (Appendix J). The Cronbach alphas of the measure of system cognitive effort, rational decision-making style, intuitive decision-making style, and perceived design technique knowledge were 0.89, 0.86, 0.88, and 0.91

[1] Before inviting experiment participants, the sample size were calculated by using G*Power 3.1.9.2 with an a priori test (Faul et al. 2007). When considering the statistical power of 0.80 for a medium effect size (f=0.25) and a significant level of 0.05 (Cohen 1988), the minimum sample size should be 159. Thus, we recruited 195 people to participate the experiment to make sure there were enough subjects in the experiment. After the experiment, five data points are removed because of incomplete or mistaken understanding of the experiment tasks.

[2] https://iism-kd2-hroot.iism.kit.edu/?locale=en

respectively, indicating adequate reliability of the measurement scales. The correlations between the three variables and selection accuracy is reported in Table 5-5.

Table 5-5. Correlations between variables (n = 190)

Variables	**Mdn**	**M**	**SD**	**Correlations**				
				1	2	3	4	5
1. Selection accuracy	8	8.59	3.20	-				
2. System cognitive effort	3	3.18	1.41	-0.47***	-			
3. Rational decision-making style	5.7	5.47	0.91	0.16*	-0.15*	-		
4. Intuitive decision-making style	3.2	3.35	1.08	-0.01	0.14+	-0.34***	-	
5. Perceived design technique knowledge	2.57	2.87	1.18	0.03	-0.04	0.06	0.15	-

*Note: +P<0.1; *P<0.05; **P<0.01; ***P<0.001*

Hypotheses Test

Two hierarchical linear regressions were used to test H1-H3 and H4a-H4b separately. After the hierarchical regression tests, further tests such as Kruskal-Wallis and ANOVA test were followed for in-depth analysis. The detail analyses and results are presented in the following.

The Effect of Decision Aids on System Cognitive Effort and Selection Accuracy

A summary of testing H1-H3 is presented in Table 5-6. Before looking into the differences between the two decision aids, a hierarchical linear regression was conducted to test whether decision aids can influence system cognitive effort and selection accuracy (Table 5-7).

Table 5-6. A summary of used tests and the results of H1-H3

#	**Hypothesis**	**Test**	**Result**
H1:	The decision aid taxonomy will lead to lower system cognitive effort than decision aid tags	Linear regression + Kruskal-Wallis Test + Wilcoxon Rank-sum Test	Supported
H2:	The decision aid taxonomy will lead to a higher selection accuracy than decision aid tags	Linear regression + Kruskal-Wallis Test + Wilcoxon Rank-sum Test	Supported
H3:	A reduction of system cognitive effort will lead to an increase in selection accuracy	Linear regression	Supported

Table 5-7. Hierarchical linear regression test relates to H1-H3 (n = 190)

	Dependent variable				
	System cognitive effort		***Selection accuracy***		
Model	Step 1: Controls-only	Step 2: Effect of decision aids (H1)	Step 3: Controls-only	Step 4: Effect of decision aids (H2)	Step 5: Effect of decision aid and cognitive effort (H3)
Age	-0.002 (0.030)	-0.004 (0.026)	0.009 (0.070)	0.012 (0.050)	0.010 (0.049)
Gender	0.253 (0.228)	0.157 (0.196)	-0.325 (0.521)	-0.020 (0.372)	0.037 (0.367)
Education	0.120 (0.214)	0.164 (0.184)	-0.095 (0.490)	-0.203 (0.350)	-0.143 (0.345)
Perceived design technique knowledge	-0.040 (0.089)	0.015 (0.077)	0.065 (0.204)	-0.094 (0.146)	-0.088 (0.144)
TAGS		**-0.805*** (0.217)**		**0.748^{+} (0.411)**	0.452 (0.420)
TAXO		**-1.777*** (0.216)**		**5.111*** (0.410)**	**4.457*** (0.472)**
SCE					**-0.368** (0.138)**
Constant	2.843*** (0.716)	3.655*** (0.626)	8.741*** (1.636)	6.941*** (1.189)	8.285*** (1.274)
R2	0.011	**0.279**	0.003	**0.499**	**0.517**
ΔR2	**-**	**0.268**	**-**	**0.496**	**0.514**
F Statistic	0.522 (df = 4; 185)	11.788*** (df = 6; 183)	0.152 (df = 4; 185)	30.327*** (df = 6; 183)	27.871*** (df = 7; 182)

*Notes: 1) the table gives coefficients (standardized errors); 2) +P<0.1; *P<0.05; **P<0.01; ***P<0.001; 3) ΔR2 scores compare the model against the controls only model; TAGS: decision aid tags; TAXO: decision aid taxonomy; SCE: system cognitive effort.*

The regression models (Step 1 and 3) of the regression of all control variables on system cognitive effort and selection accuracy showed that all control variables were not significantly regressed onto the dependent variables. When adding decision aids (taxonomy and tags) as independent variables, the R^2 of the regression model largely increased with a ΔR^2 of *0.268* (Step 2) and *0.496* (Step 4) separately. When adding system cognitive effort as an independent variable to test whether it negatively influenced the selection accuracy, the R^2 increased to *0.517* (Step 5). But the direct effect of tags on selection accuracy disappeared *($p = 0.282$)*. The robust test of the regression models (Step 2, 4, and 5) can be found in Appendix K.

Differences between Decision Aid Taxonomy and Tags (H1 and H2)

In order to have an in-depth analysis of the influence of decision aid taxonomy and tags on system cognitive effort and selection accuracy (H1 and H2), Kruskal-Wallis test and Wilcoxon Rank-sum test were used because the data of the dependent variable system cognitive effort and selection accuracy were not normal based on the Shapiro-Wilk test *($p < 0.05$)*. The boxplots and the *p-value* of the Kruskal-Wallis test and Wilcoxon Rank-sum test on the effect of decision aids on system cognitive effort and selection accuracy are presented in Figure 5-12.

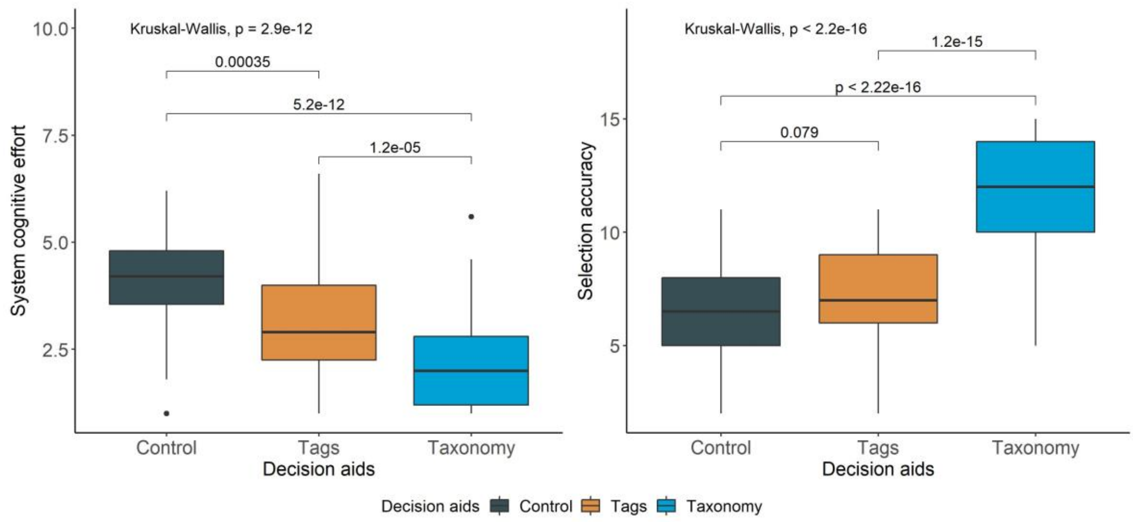


Figure 5-12. The effect of decision aids on cognitive effort (left) and accuracy (right)

The left boxplot in Figure 5-12 presents that there was a significant effect of decision aids on system cognitive effort at the $p < 0.000$ level for the three conditions $[H\ (2) = 53.11,\ p < 0.000]$. The comparisons of the mean ranks between groups showed that system cognitive effort was significantly different when a novice used decision aid tags *(difference = 31.43)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.46*. The Wilcoxon Rank-sum test showed that system cognitive effort was significantly lower with decision aid tags *(Mdn = 2.9)* than control *(Mdn = 4.2), p < 0.000, r = -0.32*. When decision aid taxonomy was used, system cognitive effort was significantly lower than control *(difference = 70.61)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.27*. The system cognitive effort was significantly lower with decision aid taxonomy *(Mdn = 2.0)* than control *(Mdn = 4.2), p < 0.000, r = -0.61*. There is a significant difference between the use of taxonomy and tags *(difference = 39.18)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.46*. The system cognitive effort was significantly lower with decision aid taxonomy *(Mdn = 2.0)* than tags *(Mdn = 2.9), p < 0.000, r = -0.39*. Hence, H1 was supported.

The right boxplot in Figure 5-12 presents that there was a significant effect of decision aids on selection accuracy at the $p < 0.000$ level for the three conditions $[H\ (2) = 90.96,\ p < 0.000]$. Subsequently, a post hoc test on the comparisons of the mean ranks between groups showed that accuracy was not significantly different when a novice used decision aid tags *(difference = 12.17)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.46*. The Wilcoxon Rank-sum test showed that selection accuracy was fairly significantly lower with no decision aid *(Mdn = 6.5)* than decision aid tags *(Mdn = 7.0), p = 0.079, r = -0.16*. When the decision aid taxonomy was used, accuracy was significantly higher than in the case of tags *(difference = 85.37)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.27*. The selection accuracy was significantly lower with no decision aid *(Mdn = 6.5)* than decision aid taxonomy *(Mdn = 12.0), p < 0.000, r = -0.74*. There was a significant difference between the use of decision aid taxonomy and tags *(difference = 73.20)* with the critical difference *(α = 0.05 corrected for the number of tests)* of *23.46*. The selection accuracy was significantly lower with decision aid tags *(Mdn = 7.0)* than with taxonomy *(Mdn = 12.0), p < 0.000, r = -0.71*. Hence, H2 was supported.

The Role of System Cognitive Effort (H3)

The hierarchical regression test (Table 5) reflects that *H3* was supported. Further analysis of the role of system cognitive effort was conducted. The mediating effect of system cognitive effort was tested based on Tingley et al. (Tingley et al. 2014). The quasi-Bayesian Monte Carlo method was used to compute the *average causal mediating effect (ACME), average direct effects (ADE)*, and *p-values*. The results showed there was a partial mediating effect of system cognitive effort between the relation of decision aid taxonomy and selection accuracy *(ACME = 0.661, p = 0.03, ADE = 4.417, p < 0.000)*. Whereas there was a full mediating effect of system cognitive effort between the relation of decision aid tags and selection accuracy *(ACME = 0.298, p = 0.014, ADE = 0.432, p = 0.326).*

Moderating Effect of Decision-making Styles

As H4a and H4b need the test of moderating effect, a hierarchical linear regression analysis was used, which was further tested by ANOVA to see whether there were differences between the model with or without the moderator (Table 5-8) (Baron and Kenny 1986).

Table 5-8. A summary of used tests and the results of H4a and H4b

#	Hypothesis	Test	Result
H4a:	Rational decision-making style moderates the relation between decision aid taxonomy and accuracy	Linear regression + ANOVA	Not supported
H4b:	Intuitive decision-making style moderates the relation between decision aid tags and accuracy	Linear regression + ANOVA	Not supported

Before analyzing the effect of rational and intuitive decision-making style on the relationship between using decision aids and selection accuracy, Kruskal-Wallis test was used to test whether there was no difference in the distribution of rational and intuitive decision-making style across the three experiment groups (taxonomy, tags, and control). The result presented that there was no difference in the distribution of rational and intuitive decision-making styles across the three experiment groups with a *p-value* of *0.621* and *0.498* respectively. The *post hoc* analysis of the Kruskal-Wallis test also showed that there was no significant difference between any two of the experiment groups.

The analysis of the moderating effect of decision-making styles was based on Baron and Kenny (1986), and a hierarchical regression was conducted. The result of the hierarchical linear regression test is presented in Table 5-9.

Table 5-9. Hierarchical linear regression for testing H4a and H4b (n = 190)

	Dependent variable: *selection accuracy*			
Model	Step1: Effect of classifications	Step 2: Rational style as a moderator (H4a)	Step 3: Intuitive style as a moderator (H4b)	Step 4: Rational and intuitive style as moderators
Age	0.012 (0.050)	0.026 (0.050)	0.010 (0.050)	0.025 (0.050)
Gender	-0.020 (0.372)	-0.033 (0.369)	0.059 (0.380)	-0.007 (0.380)
Education	-0.203 (0.350)	-0.296 (0.347)	-0.206 (0.352)	-0.283 (0.350)
Design technique knowledge	-0.094 (0.146)	-0.091 (0.145)	-0.066 (0.149)	-0.074 (0.149)
TAGS	**0.748+ (0.411)**	1.415 (2.594)	0.910 (1.326)	2.366 (3.460)
TAXO	**5.111*** (0.410)**	**9.397*** (2.417)**	**4.785*** (1.331)**	**10.498** (3.328)**
RDS		**0.624* (0.272)**		**0.630* (0.296)**
TAGS x RDS		-0.133 (0.472)		-0.205 (0.505)
TAXO x RDS		**-0.793+ (0.434)**		**-0.876+ (0.468)**
IDS			-0.207 (0.259)	0.020 (0.278)
TAGS x IDS			-0.051 (0.384)	-0.171 (0.406)
TAXO x IDS			0.104 (0.375)	-0.186 (0.398)
Constant	6.941*** (1.189)	3.401+ (1.933)	7.491*** (1.431)	3.228 (2.445)
R2	**0.499**	**0.518**	**0.503**	**0.519**
ΔR²	**0.496**	**0.515**	**0.500**	**0.516**
F Statistic	30.327*** (df = 6; 183)	21.467*** (df = 9; 180)	20.228*** (df = 9; 180)	15.932*** (df = 12; 177)

*Notes: 1) the table gives coefficients (standardized errors); 2) +P<0.1; *P<0.05; **P<0.01; ***P<0.001; 3) ΔR^2 scores compare the model against the controls only model; 4) RDS: rational decision-making style; IDS: intuitive decision-making style; TAGS: decision aid tags; TAXO: decision aid taxonomy.*

The regression of all control variables and decision aids onto the selection accuracy (Step 1) is the model without any moderating effects. For testing H4a, the moderating effect of rational decision-making style was added. The regression model in Step 2 (H4a) presented there was a significant direct effect of ration decision-making style on selection accuracy *(p = 0.023)* and a slightly significant effect of the interaction of rational decision-making style and decision aid taxonomy *($p = 0.069$)*. The R^2 of the model without the moderator and with the moderator increased from *0.499* to *0.518*. Robust test result of H4a is in Appendix K. An ANOVA test was used to compare the two models in Step 1 and 2. The results showed that the regression model with the moderating effect of rational decision-making style modestly significantly improved the fit of the model than without the moderator *($p = 0.071$)*. Hence, there was a moderating effect of ration decision-making style between the relationship of decision aid taxonomy and selection accuracy. But the coefficient of the interaction between rational decision-making style and high structured classification was negative, which indicated that rational decision-making style negatively moderated the influence of decision aid taxonomy towards selection accuracy. The result was opposite to the hypothesis. Hence, H4a was not supported. When looking at the test of the moderating effect of intuitive decision-making style (H4b), the regression model (Step 3) presented that there was no direct effect of intuitive decision-making style or interaction effect with any decision aids on the selection accuracy. Additionally, when adding both rational and intuitive decision-making styles to the regression model (Step 4), the R^2 nearly did not change when comparing with the model for testing H4a. Hence, H4b was not supported.

Figure 5-13 shows a deeper analysis of H4a. The regression lines of the interactive effect of rational decision-making style and decision aids (left in Figure 5-13) reflects that with the increase of rational decision-making style, there was a decrease in selection accuracy when using decision aid taxonomy. The simple slopes for selection accuracy were divided at three levels of rational and intuitive decision-making styles (1SD above the mean, mean, and 1SD below the mean) (Aguinis and Gottfredson 2010; cf. Artz et al. 2012) (right in Figure 5-13). The standard deviation analysis of rational decision-making style reflected that when providing a taxonomy, high rational decision-makers had a lower selection accuracy than low rational decision-makers. For low rational decision-makers (-1SD in Figure 5-13), the effect of decision aid taxonomy *($beta = 1.23$, $t = 0.30$, $p < 0.000$)* was positive, but the interaction effect of rational decision-making style was negative on the selection accuracy, with *($beta = -0.54$, $t = 0.31$, $p = 0.086$)*. For high rational decision-makers (+1SD in Figure 5-13), the effect decision aid taxonomy *($beta = 1.43$, $t = 0.42$, $p < 0.001$)* was positive, but the interaction effect of rational decision-making style was negative on the selection accuracy, with *($beta = -0.73$, $t = 0.42$, $p = 0.086$)*. When comparing the regression line of high rational decision-makers and low decision-makers, the two standard coefficient values showed that the higher the rational decision-making style, the lower the selection accuracy.

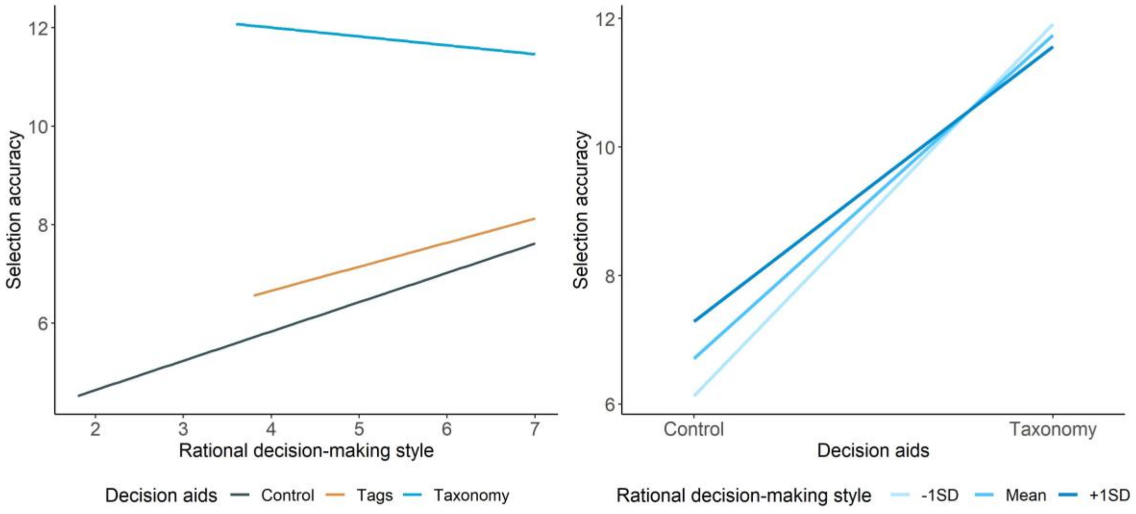


Figure 5-13. Linear regression analysis (left) and standard deviation analysis (right)

5.2.5 Discussion of Results

The experiment results showed that the taxonomy could not only directly improve the selection accuracy but also improve it by reducing the system cognitive effort spent on understanding the unfamiliar information. Furthermore, the rational decision-making style presented a certain trend toward significantly moderated the relationship between the use of decision aid taxonomy and selection accuracy. The regression lines in Figure 5-13 presented that the relationship between the use of high structured classification and selection accuracy was moderated by rational decision-making style. The regression lines presented different trends of the selection accuracy with the increase of rational decision-making style. When using the decision aid taxonomy, the selection accuracy declined slightly with the increase of rational decision-making style. Whereas, in the decision aid tags and control group, the selection accuracy increased with the rise of rational decision-making style. The results implied that the structure of categories of design techniques in the decision aid taxonomy did not totally match the structure of categories of design techniques in the mind of experiment participants with high rational decision-making style. As the experiment participants were design novices, although the rational decision-makers tend to get to know all the information of design techniques before making the decision, the limited knowledge and expertise negatively influenced the decision-making process. Whereas, in the decision aid tags and control group, the higher the rational decision-making style, the better the selection accuracy. Such a result reflected that novices can benefit from rational decision-making style in the selection of design techniques when there was no decision aid or with decision aid tags. Compared with the three regression lines, the regression line of decision aid taxonomy was always above the regression line of tags and control group (left in Figure 5-13). The differences presented that although the structure of design techniques in the decision aid taxonomy was not completely consistent with the understanding of novices, the use of taxonomy could still lead to a better decision outcome.

Compared with the decision aid taxonomy, the positive effect of tags on selection accuracy showed a quasi-significant. But there was a full mediating effect of system cognitive effort in the relationship between decision aid tags and selection accuracy. The negative effect of decision aid tags on system cognitive effort implied that the categories with no parent-child relations could also help with understanding the categories when selecting design techniques. However, there was no moderating effect of intuitive decision-making style between decision aid tags and selection accuracy. Although the holistic decision-making style leads to the creation of flat categories (Goh et al. 2009; Golder and Huberman 2006; Klašnja-Milićević et al. 2017), the match in between did not positively influence the selection accuracy in our experiment. One of the reason could be that when selecting design techniques, the intuitive decision-makers followed their feelings and hunch which could be different from person to person and also different from the flat categories included in the decision aid tags (Epstein et al. 1996; Hamilton et al. 2016).

6 Design Cycle 3: Advisory-based Design Support System with Embedded Taxonomy[1]

Based on the result of design cycle 2, compared with tags, taxonomy led to higher selection accuracy and lower cognitive effort. In cycle 2, the taxonomy was visualized as a navigation menu in the web-based platform and in the experiment tool. The dimensions and characteristics of the taxonomy were abstract and not self-explained. Users needed to map the design situations in their mind with the abstract terms in the taxonomy. If the dimensions and characteristics were explained in natural language, the abstract terms would be directly explained, which could reduce the complexity caused by understanding the abstract terms. Thus, in cycle 3, an iterated version of the web-based platform which attempted to use natural language dialog with embedded taxonomy to suggest design techniques was provided. A live system was developed, and interviews, as well as questionnaire, were used to test the usability of the platform. In order to evaluate whether the different UIs could influence the selection performance and whether different decision-making styles could moderate the relation between using different UIs of taxonomy and the selection performance, a lab experiment was conducted. Figure 6-1 presents the overall research approach in design cycle 3. Part 1 describes the development of the iterated web-platform (ServiceDesignKIT 2.0), while part 2 describes the experiment design and the hypotheses test about the different effects of two UIs of taxonomy on the selection performance with the consideration of decision-making styles.

	Part 1 (Chapter 6.1)	Part 2 (Chapter 6.2)
Instantiation and artifact	An iterated version of the web-based platform with an attempt to build an advisory platform with the embedded taxonomy	Develop a system including a taxonomy UI and a natural language UI for an experiment
Evaluation	Usability tests	Lab experiment
Result	Design principles for developing an advisory platform for supporting the selection of design techniques	Explain how different UIs influence the selection of design techniques

Figure 6-1. Research approach for cycle 3

[1] This Chapter is based on the following studies which are published or in work: Liu, He, et al. (2019); Liu, Rietz, et al. (2019).

6.1 Designing an Advisory Platform for Supporting the Selection of Design Techniques (ServiceDesignKIT 2.0)

In this Chapter, ServiceDesignKIT 1.0 and 2.0 are compared. Specifically, the derived meta-requirements are introduced, the design principles for addressing the meta-requirements are proposed, and the design features for instantiating the design principles are presented. The instantiation as well as the result of a usability test are explained.

6.1.1 A Comparison between ServiceDesignKIT 1.0 and 2.0

A summary of the comparison between ServiceDesignKIT 1.0 and 2.0 is introduced before introducing meta-requirements and design principles for ServiceDesignKIT 2.0 in detail. Figure 6-2 presents a comparison of meta-requirements. For MR1, in ServiceDesignKIT 2.0, the form of the decision aid for helping the selection does not limit to classifications. The platform can be extended by using other forms of decision aids. MR2 in ServiceDesignKIT 2.0 combines MR2-MR4 from ServiceDesignKIT 1.0. The requirements related to enabling users to suggest design techniques, add comments, and save favorite techniques are integrated into one. Moreover, in ServiceDesignKIT 2.0, users should be enabled to organize the stored design techniques into specific sets based on different design projects or design situation. MR3 is a new meta-requirement that suggests providing a new form of decision aid that can guide design novices in the selection process.

ServiceDesignKIT 1.0	ServiceDesignKIT 2.0	Comparison
MR1. Provide a clearly structured classification as a filter to support searching and finding suitable design techniques.	MR1. When users need to find design techniques, the system should provide a design technique library based on different dimensions and characteristics.	In ServiceDesignKIT 2.0, the form of decision aid for selection design techniques does not restrict to classifications.
MR2. Enable users to extend and improve the contained design techniques. MR3. Enable users of the web platform to comment and discuss design techniques. MR4. Enable users to save their favorite design techniques to a shortlist.	MR2. When users find appropriate techniques, the system should provide users a dedicated area to store them. Additionally, the system should allow users to submit new design techniques.	The requirements related to individual space (ServiceDesignKIT 1.0) are integrated into one (ServiceDesignKIT 2.0).
	MR3. For novice users with less experience in design processes, the system should guide the selection of design techniques in a step-by-step approach	The system can guide the selection process of design techniques.

Figure 6-2. Comparison of meta-requirements

Figure 6-3 provides a comparison of design principles. For DP1, in ServiceDesignKIT 2.0, taxonomy is used as a main decision aid, while tags are used as a supplement. Because taxonomy is proved that it leads to higher selection accuracy in the lab experiment in part 2 of cycle 2 (Chapter 5.2). For DP2, the design feature for making comments on each design technique is moved to DP4 which enables to make comments for a group of design techniques for a specific design project. DP3 and DP4 are new design principles in ServiceDesignKIT 2.0, which suggests providing a natural language UI for step-by-step guiding design novices' selection process. In addition, the selected design techniques by using the natural language UI should be grouped into specific design project and be able to be stored in the individual space.

ServiceDesignKIT 1.0	ServiceDesignKIT 2.0	Comparison
DP1. Provide classifications (taxonomy and tags) and controlled users' suggestion of design techniques	DP1. provide taxonomy as a main decision aid and tags as supplements to help users to select design techniques	Taxonomy is main selection support. Taxonomy and tags are not combined, tags are supplements to the taxonomy.
DP2. Include communication features and personalization features.	DP2. Provide the system with a personal page to store the favorite design techniques and the suggestion history	The feature for making comments is moved to DP4 which enables to comment on a group of techniques for a specific project.
	DP3. provide a natural language UI to suggest design techniques DP4. provide space for saving suggestion design techniques based on different project	Compared with the abstract taxonomy, design techniques are suggested by using natural language dialog. And the suggestion history is saved and sorted by different design projects.

Figure 6-3. Comparison of design principles

6.1.2 Research Approach

Similar to the research approach for developing ServiceDesignKIT 1.0, Peffers et al (2007)'s method was followed. Three meta-requirements focusing on supporting the selection of design techniques and instructing the selection process were derived based on problem analysis in selecting and suggesting design techniques for different design situations. In order to address the meta-requirements, four refined design principles were proposed to address the meta-requirements focusing on suggesting and organizing design techniques. Subsequently, design features that addressed the design principles were suggested and implemented, e.g., using a natural language UI with embedded taxonomy to guide the selection of design techniques. The developed prototype was evaluated by in-depth interviews (Myers 1997) and the system usability scale (SUS) questionnaire (Brooke 1996).

6.1.3 Problem Awareness and Meta-Requirements

The first meta-requirement (MR1) refers to offering users an organized design technique library with information on different dimensions and characteristics of design techniques. As the collection of design techniques is enriched by the absorption of techniques from other disciplines, the number of design techniques expands (Moritz 2005).

With a large number of design techniques, choosing the appropriate ones for a specific design situation is a challenge (Fuge et al., 2014; Sanders et al., 2010; Sharma et al., 2017). Face-to-face consultancy is a direct way to get suggestions for appropriate design techniques for specific design situation. However, consultancy service is costly. Alternatively, design novices can use a search engine with keywords to acquire design techniques, which has many weaknesses. The search process by using keywords usually needs a large amount of time and may result in low accurate information because of the huge volume of information and irrelevance of keywords (Iqbal et al. 2017). People need to invest a lot of effort to filter and synthesize the desired information. As a clear overview of different design techniques with dimensions and characteristics can present the differences and similarities and narrow down the filtering scope, it is important to build a library for design techniques. This should enable users to distinguish and match design techniques with different design situations.

MR1: *When users need to find design techniques, the system should provide a design technique library based on different dimensions and characteristics.*

The second meta-requirement (MR2) refers to the customization of the suggested set of design techniques. Personalization is known to influence the usage of information systems (Kim and Son 2009). When working on a design project, different techniques are needed in different phases of the design process. People have their own preferred and frequently used design techniques. If a group of techniques can be organized with the corresponding design project and recorded as search history, design novices can build their own technique sets. A UI for each design novice to access the frequently used techniques by saving them along with the suggestion history should be provided. This personalization feature allows novices to tailor specific sets of design techniques for specific projects or design situations. Thus, a dedicated area for storing suggested design techniques should be provided. In addition, users should be enabled to submit new design techniques or give advice to the existing design techniques in the system.

MR2: *When users find appropriate techniques, the system should provide a dedicated area to store the suggested results. Additionally, the system should allow users to submit new design techniques.*

The third meta-requirement (MR3) refers to providing guidance to design novices when selecting design techniques. When comparing with design experts, design novices are characterized by a different cognitive process and viewpoints of design processes. For example, when designing a webpage, design novices need more specific instructions such as fully described requirements for conducting design activities (Bonnardel et al. 2003). A reason for this result is the lack of professional knowledge. A detailed and concrete description of design situations can help design novices in planning design activities. Furthermore, as the design activities may not be the main task or are occasional or infrequent for design novices, they may not invest as much time as design experts in studying design knowledge or training design practices (Bonnardel et al. 2003). Thus, if the platform can help design novices in constructing and describing their requirements clearly, they will gain a deeper insight into the required design techniques. Therefore, step-by-step instructions for design novices during the selection process need to be provided.

MR3: *For novice users with less experience in design processes, the system should guide the selection of design techniques in a step-by-step approach.*

6.1.4 Design Principles

In order to address the derived meta-requirements, several design principles (DPs) were suggested. The proposed design principles do not map one-to-one to the meta-requirements. A meta-requirement can be addressed by multiple design principles, and vice versa.

As design techniques can be selected based on their different characteristics and design situations (MR1), the system needs to include different characteristics of design techniques. As classifications can reduce the complexity and enable the systematical comparison of data, using classifications comes with cognitive benefits (Gorlenko and Englefied 2006; Parsons and Wand 2008). Classification is widely used in information system development, such as objected-oriented system design, data management, and ontological engineering (Parsons and Wand 2008). The process of using classification is also a learning process, and it helps people to understand how information is organized and sought out (Roschuni et al., 2015). With a classification of design techniques, people can decide which design techniques are most relevant to a specific design situation (Sanders et al. 2010). As forms of classifications, top-down taxonomy (Liu et al. 2016) (part 1 of cycle 1, Chapter 4.1) and bottom-up tags (part 2 of cycle 1, Chapter 4.2) were used as a filter in ServiceDesignKIT 1.0 (part 1 in cycle 2, Chapter 5.1). However, the lab experiment in part 2 of cycle 2 (Chapter 5.2) indicated the taxonomy outperformed tags in selection accuracy. Thus, for ServiceDesignKIT 2.0, the taxonomy can be the main decision aid and tags can be supplements. For each design technique, the basic information (i.e., name, description, instruction, and reference) and the corresponding attributes (i.e., characteristics and dimensions in the top-down taxonomy as well as bottom-up tags) should be stored in the database. The collection of techniques should a basic library for the application.

DP1: *Provide the system with a top-down taxonomy as a main decision aid and bottom-up suggested tags as supplements for users to select design techniques.*

MR2 emphasizes that design novices should be encouraged to participate in design activities. As an emerging field, service design adopts techniques and knowledge from different disciplines, which enriches the design technique library for different design activities (Mager 2009). People with different knowledge backgrounds should have the possibility to engage in digital service design practices. Some web-based applications are developed aiming to build a design community such as the web portal *theDesignExchange* focusing on sharing design knowledge and experiences (Roschuni et al. 2011). Thus, ServiceDesignKIT 2.0 should provide the function enabling users to suggest new design techniques to the technique library and assign design techniques to the appropriate dimensions, characteristics, and tags. This enables the system to directly associate a design technique with the classifications of the design technique library. However, the suggested techniques may be varied in content and quality, which influences the quality of the the platform (Xu et al. 2011). Thus, the suggested design techniques should be at first stored in a temporary database, which enables the administrator to review and edit them. And then, those with high quality will be added into the technique library.

DP2: *Provide the system with a function for users to suggest new design techniques with dimensions and characteristics to the technique library.*

As the system should specifically instruct design novices when selecting design techniques (MR2), step-by-step suggestions need to be given when users are selecting design techniques. As natural language dialogs are easy to understand and provide an intuitive way to access information, a natural language UI can be used to guide the selection process. The concept has been applied in many fields such as filtering information, retrieving information, managing emails, scheduling meetings, etc. (Allison 2012). A natural language UI can bring novices benefits by reducing the information overload and offering an intuitive insight into the problem (Sharma et al. 2017). With this intuitive conversation-like approach (i.e., a process of questions and answers), design novices should be able to articulate their needs clearer and better consider critical factors of choosing digital service design techniques (Carayannopoulos 2018). Thus, a well-structured natural language UI should be provided.

DP3: *Provide the system with a natural language UI to guide design novices in the selection process of design techniques.*

As the advisory platform should have a dedicated UI to record the preferred design techniques and the suggestion history of the natural language UI (MR3), a personal page should be provided. When users find some techniques which meet their objectives, they may consider saving them for future retrieval. In many web-based applications, the personal page is expected to improve the user experience (Dumais et al. 2014). Thus, a favorite list should be provided to each user to store the design techniques which they like. Moreover, in a particular design situation, usually, a design technique is used with a combination of other design techniques (Woolrych et al. 2011). Thus, users should be enabled to organize the saved design techniques based on design projects or design situations. Besides, the suggestion history of design techniques by the natural language UI can be recorded in a timeline which includes the project information and the suggested design techniques. Users can view such information to track the suggestion history for specific design projects and design situations.

DP4: *Provide the system with a personal page for users to store design techniques based on different design projects and design situations.*

The four design principles are proposed for addressing the derived meta-requirements. Table 6-1 presents a short summary of design principles and the associated meta-requirements.

Table 6-1. Design principles and the associated meta-requirements for ServiceDesignKIT 2.0

DPs	Description	Addressed MRs
DP1	Provide taxonomy as a main decision aid and tags as supplements to help users to select design techniques	MR1, MR2
DP2	Provide the system with a personal page to store the favorite design techniques and the suggestion history	MR2
DP3	Provide a natural UI to suggest design techniques	MR3
DP4	Provide space for saving suggested design techniques based on different projects	MR2

6.1.5 Prototypical Implementation

In order to develop a prototype of the system, design features (DFs) are provided based on the meta-requirements and design principles. Design features are specific capabilities of the system to satisfy design principles (Meth et al. 2015).

To satisfy DP1, a design technique card gallery with a filter can be used to visualize the classified design techniques. Each design technique is presented as a card which contains the technique name and an image. The filter is constructed based on the dimensions and characteristics of a taxonomy (Liu et al. 2016) (part 1 in cycle 1, Chapter 4.1) and bottom-up suggested tags (part 2 in cycle 1, Chapter 4.2) by design novices.

DF1: *Use a card gallery to visually depict design techniques with the combination of a taxonomy and tags to enable filtering*

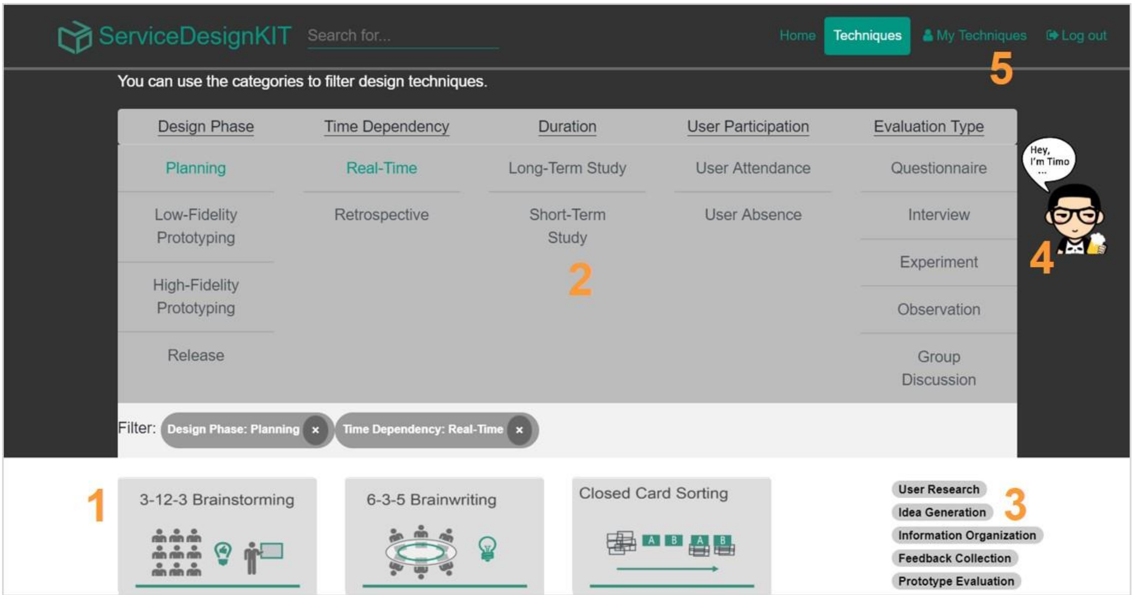


Figure 6-4. A card gallery of design techniques with a filter (servicedesignkit.org)

Figure 6-4 presents a screenshot of the advisory platform based on DF1. Design techniques are visualized as a set of cards (1 in Figure 6-4). Users can click on each design technique card to open a webpage which includes detailed information about the design technique. The characteristics under each dimension can be selected exclusively, and all the five dimensions can be selected collectively (2 in Figure 6-4). On the right side, a tag cloud is depicted (3 in Figure 6-4), which includes bottom-up suggested tags from design novices. Users can also filter design techniques by using these tags. Each time the user changes the selected characteristics or tags, the cards gallery will be refreshed and show corresponding design technique cards. In order to instruct the selection process, the natural language UI (4 in Figure 6-4) shows up at the same position on each webpage. Details on the natural language UI are described in DF2.

To satisfy DP2, a natural language UI is provided to interact with users and guide them to find appropriate design techniques. The UI is designed as a chat frame.

DF2: *Embed a natural language UI in order to guide users to find appropriate design techniques for specific design situations*

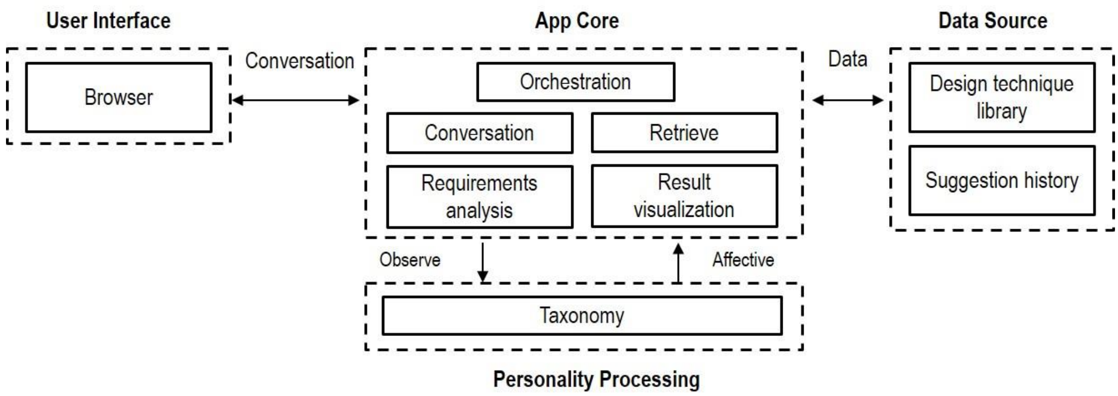


Figure 6-5. The architecture of the natural language UI, based on Di Prospero et al. (2017)

Di Prospero et al. (2017) suggest an architecture framework for conversational agents. This architecture divides a conversational agent into four important components: user interface, application core, external data sources, and personality processing. The four components are applied to identify the architecture of the natural language dialog (Figure 6-5). To make access convenient and easy for users, the natural language UI interacts with users based on the browser. To provide design novices appropriate design techniques, a technique library, which includes the necessary information of techniques is associated as an external data source. The suggested techniques will be stored in the suggestion history for users to retrieve. The application core is the component which monitors and controls the conversation (Di Prospero et al. 2017). After the result of recommendation arrives, the application core visualizes the result and presents to users. In personality processing, the natural language UI matches the needs of users with the taxonomy of design techniques to suggest suitable design techniques to users.

In order to make sure the user requests match the database, a pre-defined conversation structure needs to be applied in the natural language UI to guide the process of conversation (Allison 2012). The conversation should guide users to the topics relevant to the suggestion of design techniques.

The conversation logic is presented in Figure 6-6. The natural language UI can guide users step-by-step to the two main functions: searching or recommending design techniques. If a user chooses "Let me recommend," the natural language UI will at first present several pre-defined questions on the user's design situations, and then map the answers to the dimensions and characteristics in the taxonomy to calculate the similarities to recommend design techniques.

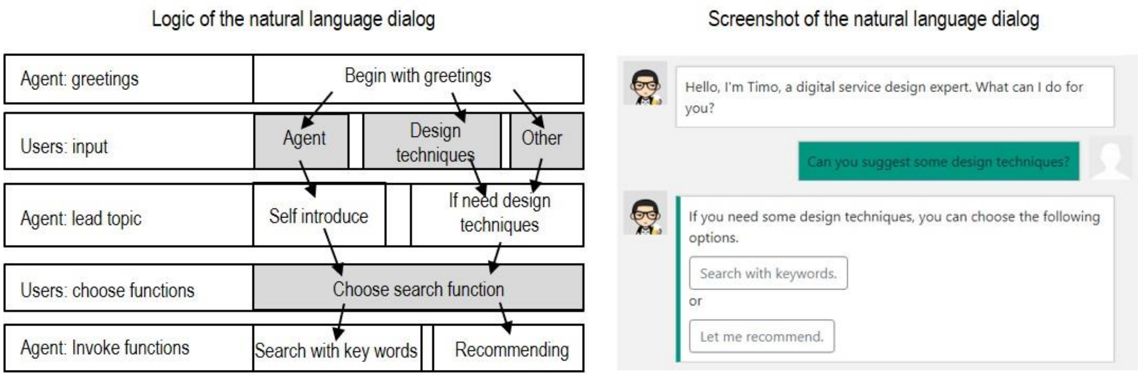


Figure 6-6. The logic of the natural language UI (left) and a screenshot (right)

Finally, to satisfy DP3 and DP4, a personal page should provide the function of saving a favorite list as well as the suggestion history from the natural language UI, and a form to submit new design techniques. These functions will be associated with a personal page including different panes which can be switched by using a navigation list.

DF3: *Enable users to save favorite design techniques, to capture the history of suggested design techniques, and to submit new design techniques by providing a personal page*

By clicking on "My techniques" at the top right of the navigation bar (5 in Figure 6-4), the individual view will appear (Figure 6-7). The individual view includes three panes (1 in Figure 6-7): a favorite list (i.e., My techniques), suggestion history (i.e., Advisory History), and submission of new techniques (i.e., Suggest New Techniques). The favorite list stores the technique cards which are marked as liked by users. In the suggestion history, the suggested design techniques by the conversational agent are stored in a timeline and grouped by projects with the project name as well as design situations (2 in Figure 6-7). The design situations are used as selection conditions to match the dimensions and characteristics in the taxonomy. Users are allowed to add comments to the suggested design techniques (3 in Figure 6-7). In addition, a feedback question is added next to the suggested design techniques (4 in Figure 6-7). Knowing whether the suggestion is helpful for users for a specific design project can help with the further improvement of the web-based platform. In the pane of submission of new techniques, there is a form enabling users to submit new design techniques to the system for enriching the database.

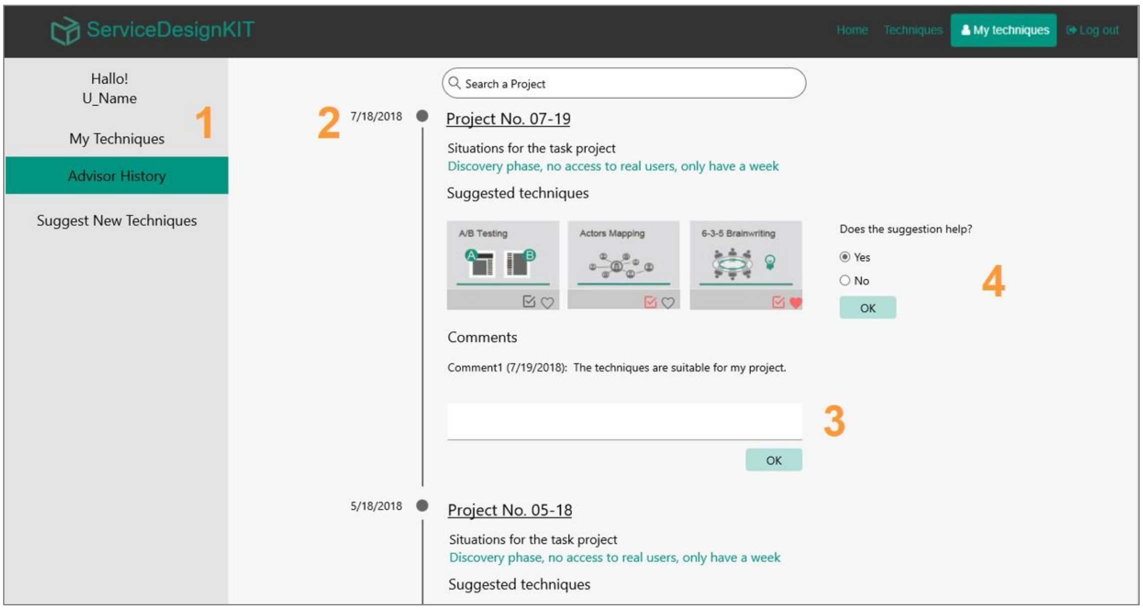


Figure 6-7. Individual space of ServiceDeisgnKIT 2.0

Table 6-2 presents a summary of design features and addressed design principles.

Table 6-2. Design features and associated design principles for ServiceDesignKIT 2.0

DFs	Description	Addressed DPs
DF1	A card gallery of design techniques with taxonomy and tags to enable filtering	DP1
DF2	A natural language UI to guide users to find appropriate design techniques	DP3
DF3	A personal page to save favorite design techniques, add comments, and enable the submission of new design techniques	DP2, DP4

6.1.6 Usability Evaluation

To evaluate the design principles and the instantiated advisory platform, a semi-structured interview (Myers 1997) with a SUS questionnaire (Brooke 1996) were combined. The semi-structured interview focuses on collecting in-depth feedback from the participants. The SUS questionnaire seeks to measure the perceived usability with a ten-item scale (Brooke 1996). The five-point scale ranging from “strongly agree” to “strongly disagree” was used (Brooke 1996). Each participant was asked to rate the ten statements (i.e., items). In the analysis, the SUS questionnaire yields a single number, ranging from 0 to 100 (Brooke 1996). Eight students with education background of information systems were invited to take part in the evaluation. The students were recruited from the attendees of a colloquium in which the basic function of ServiceDesignKIT 2.0 was presented. The attendees were asked whether they were interested in ServiceDesignKIT 2.0 and whether they could participate. An appointment was set up with each of them who agreed to participate in the evaluation. The evaluation then started a week after the colloquium. The eight participants included three females and five males with an age range from 24 to 30. Three were doctoral students, three were bachelor students, and the other two were master students. Two of them evaluated themselves as experts in the digital service design field, and six of them evaluated themselves as novices. The participants were considered as an appropriate target group for this evaluation, as they could be potential users of this platform.

Figure 6-8 presents the evaluation (first four steps) and analysis approach (the last step). Each of the participants of the evaluation went through the first four steps in Figure 6-8. At first, each participant was given the same task (Appendix L). The participants were told to select design techniques based on this task scenario and save the favorite ones by using the taxonomy and tags (DF1), the natural language UI (DF2), and the individual page (DF3). After introducing the task scenario, the participant started to use the advisory platform.

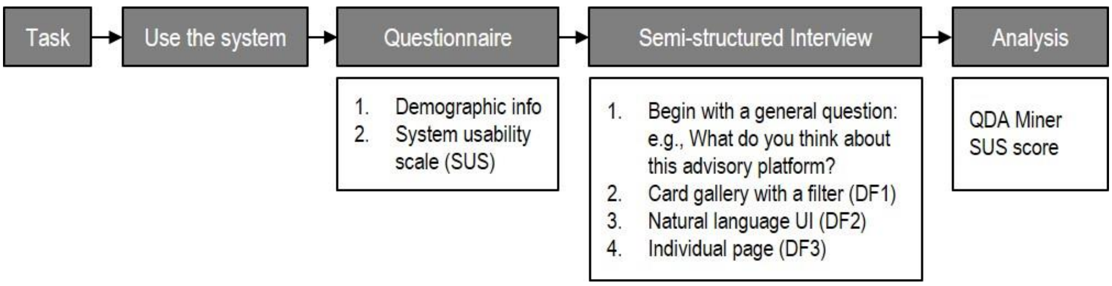


Figure 6-8. Usability evaluation and analysis approach for ServiceDesignKIT 2.0

After completing the task, at first, each participant was given a questionnaire with SUS questions and demographic questions. Subsequently, a face-to-face interview was conducted with each participant about the feedback on the system. The interview questions (Appendix L) were designed based on the provided design features of the advisory platform. Some of the questions were pre-defined for an appreciative and laddering interview (Schultze and Avital 2011). Each participant used 20-30 minutes for the task, questionnaire, and interview. Each interview was conducted face-to-face and lasted about 10-20 minutes. All interviews were recorded and transcribed for the analysis. The interview transcript was analyzed by using QDA Miner 5[1] which is a qualitative data analysis tool to achieve reliable coding of sentences or paragraphs. In order to match the feedback with the design features and evaluate each design feature, open coding was used to analyze the transcript of the interviews (Myers 1997). Table 6-3 presents two examples of the coding approach.

Table 6-3. Examples of analyzing the interview transcript

Design Feature	Coding	Quote
DF3: Conversational agent	Structured decision making	"He can advise user; it helps me with finding and structuring my decision making. […] the structure of the question helps me to categorize and know what do I really want in the situation." (P2)
	Doesn't require expertise	"It really helps me, it is very quick, you have just a few questions, and you can answer even you don't have any background knowledge." (P6)

The analysis of the SUS results from the eight participants of the evaluation presented that the average SUS score of the system was 75.31 (SD = 10.11). A system is passable when the SUS score is above 70 (Bangor et al. 2008). Furthermore, existing studies show that the total average SUS score of the web-based application is 68.2 (Bangor et al. 2009). Thus, ServiceDesignKIT 2.0 is in the acceptable range.

In the semi-structured interviews, all of the interviewees expressed their willingness to use this application in the future because the platform provided *"a collection of different kinds of techniques" (P7).* P2 and P3 shared that they hadn't seen other similar tools. P1 said it was useful because it really helped him to find design techniques. Half of the participants pointed out that the platform was well designed. Except for the generic evaluation, there was detail positive and negative feedback for each design feature. Based on the analysis of the feedback, some possibilities for optimization were suggested. Table 6-4 presents a summary of the evaluation along with the three key design features. As the participants were allowed to mention more than one positive and negative point or no feedback to each design feature, the percentage numbers in Table 6-4 don't total to 100%.

[1] https://provalisresearch.com/products/qualitative-data-analysis-software/

Table 6-4. Evaluation results with the percentage of interviewees mentioning each point

DFs	Positive feedback	Negative feedback	Suggested improvement
DF1: card gallery	• Convenient to select (50.0%) • Easy to understand (25.0%)	• Requires understanding of categories (87.5%)	• Introduction and instruction video (12.5%)
DF2: natural language UI	• Structured decision-making (62.5%) • Doesn't require expertise (62.5%)	• More serious (12.5%) • More annoying (25.0%)	• Offer using examples (37.5%) • Reduce the number of suggestion (25.0%)
DF3: personal page	• Techniques associated with projects (25.0%) • Convenient to use (87.5%)	• Need an evaluation of newly added techniques (37.5%) • Need checking the existence of suggested techniques (25.0%)	• Export the suggestion result (12.5%) • Share with group members (12.5%)

The evaluation of DF1 presented that the card gallery of design techniques with the taxonomy and tags was convenient for people to search design techniques, but the dimensions and characteristics of the taxonomy were considered not easy to understand. Here are some quotes from the interviews. P2 gave positive feedback: *"I think it is useful. [...] I can just click the options."* For supporting the understanding of the content of the taxonomy, P5 suggested: *"It will also be nice to have an explanation of what are the categories, and what is the difference between characteristics."*

The evaluation of DF2 presented that the natural language UI was well received by the participants because it could help with structuring the decision-making process, which was easy to understand. But there was still room for improvement. The natural language UI could be more intelligent. The positive feedback includes: *"it doesn't require any expertise" (P5); "It can help. Especially the agent, [...], it helps me with finding and structuring my decision-making" (P7).* P7 also mentioned the improvement, *"maybe some general question that I can ask him like not only about the conditions, like if I have some keywords that I want to ask him. He should be a little more intelligent."* Further suggestions mentioned by P2 and P6 included reducing the number of the suggested design techniques and providing examples of using the design techniques.

The evaluation of DF3 presented that the personal page was considered to be useful because participants could store and organize design techniques for specific design projects. Some improvements were also mentioned, for example, enabling the sharing of favorite design techniques and checking the existence of the suggested design techniques in the database. Here are some representative quotes. P7 mentioned the advantages, *"I think it is very good. I can see what techniques I had viewed before and liked."* The mentioned improvements are as follows. P6 mentioned, *"my suggestion is maybe you can add a different project with different techniques."* P5 suggested checking if the suggested techniques were already existing in the database. P1 talked about the sharing function, *"like sharing social network, or at least team working, some people can be added to this group, and people in this group can share examples."*

6.1.7 Conclusion

A refined version (ServiceDesignKIT 2.0) of the web-based platform (ServiceDesignKIT 1.0) was proposed, and a comparison between the design principles was provided. In order to develop the artifact, three meta-requirements were derived, and four design principles to address the meta-requirements were proposed. The proposed design principles contribute to design knowledge theoretically by guiding the development of a class of advisory systems with the specific purpose of suggesting design techniques (Gregor, 2006). In order to satisfy design principles, three design features were provided and implemented in the advisory platform. ServiceDesignKIT 2.0 was evaluated by using interviews and SUS questionnaire. Several areas of improvements were identified.

The comparison of the DPs of ServiceDesignKIT 1.0 and 2.0 presents that the most obvious difference was the implementation of the natural language UI to support the selection of design techniques, which was an attempt to visualize the taxonomy by using natural language dialog, which received positive feedback in the evaluation. But the feedback was based on the usability which did not reflect the actual performance of selection by using the natural language UI. It was obvious that the natural language UI was easier to understand than the abstract terms in the taxonomy. However, the process of suggesting design techniques by using the natural language UI was not as transparent as using a taxonomy. Because when using a natural language UI, the filtering process is not presented, and the suggested design techniques appear at the end of the conversation. Whereas, when using a taxonomy, the design techniques included by different categories will appear step-by-step by choosing one category after another — the two different UIs may influence the selection process. Hence, an experimental evaluation comparing the taxonomy and a natural language dialog should be conducted.

6.2 The Effect of Taxonomy UI and Natural Language UI on the Selection of Design Techniques

The comparison between ServiceDesignKIT 1.0 and 2.0 shows that selecting design techniques by using a taxonomy with abstract terms is different from using a natural language dialog. These two different types of UIs exist quite often on websites, including a navigation menu and a natural language-based conversational UI (Bussolon 2009; Chen et al. 2017). There are studies focusing on designing a conversational agent to guide users to find the desired information for a specific task (e.g., Allison 2012; Carayannopoulos 2018; Graf 2015; Di Prospero et al. 2017). However, to the best of my knowledge there is currently no research that compares the different effect of these two different UIs on task performance. Thus, in this Chapter, the effects of two UIs (taxonomy vs. natural language dialog) on the selection of design techniques with consideration of rational decision-making style are explained.

6.2.1 Research Approach

The research approach is presented in Figure 6-9.

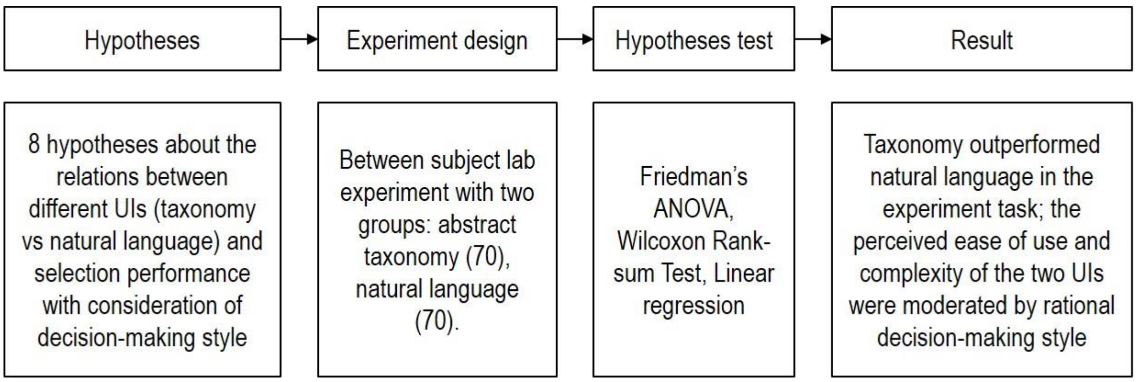


Figure 6-9. Research approach for evaluating taxonomy UI and natural language UI

In a first step, hypotheses about the differences in selection performance when using a taxonomy or natural language dialog were proposed. In order to test the hypotheses, a lab experiment was conducted. Overall, 148[1] participants (invited from KIT Hroot[2]) participated in this experiment in

[1] 140 participants provided valid data for the experiment analysis. Eight data points were removed because of incomplete, misunderstanding of the experimental tasks, or randomly choosing the answers.

[2] https://iism-kd2-hroot.iism.kit.edu/?locale=en

the KIT KD2Lab[1]. After the experiment, data was analyzed in order to test the proposed hypotheses. The result showed that natural language dialog costed less time than taxonomy for one-time use, but more time for repeated use. In general, the taxonomy outperformed the natural language dialog in selection accuracy and perceived task performance. Moreover, an in-depth analysis presented perceived complexity and perceived ease of use were moderated by the rational decision-making style.

6.2.2 Hypotheses

The hypotheses were proposed based on cognitive fit theory (Vessey and Galletta 1991)[2]. The cognitive fit model consists of the key constructs: problem representation, mental representation, and problem solution. The problem representation maps to the treatments, the taxonomy and the natural language dialog. The mental representation is covered by perceived complexity and perceived ease of use of the treatment implemented in the experimental tool (i.e., taxonomy vs. natural language dialog). Finally, the problem solution is mapped to selection accuracy and perceived task performance. In addition to the constructs included in the cognitive fit model, rational decision-making style[3] is also considered as a predictor that can influence the effect of different UIs (i.e., taxonomy UI and natural language UI) when selecting design techniques for a specific task.

Influence of UIs on Time used for each Task

It is acknowledged that expertise can be gained by repeated practices (Blum and Koskinen 1991). The more times people practicing, the more skillful people will become when performing the activity, which could contribute to an increase in efficiency. For example, research on reading skill indicates that repeated reading leads to an improvement in reading rate and comprehension (Samuels 1979). IS research also has shown that people with high experience in a specific information technology (i.e., years of experience) use less time and have a higher efficiency than people with low experience (Collins et al. 2008; Mennecke et al. 2006). As this research focuses

[1] https://www.kd2lab.kit.edu/english/21.php

[2] In part 2 of cycle 2 (Chapter 5.2), besides the external problem representation, the internal problem representation (Shaft and Vessey 2006) (perceived design technique knowledge) was also considered. But as the experiment participants were design novices and the average perceived design technique knowledge was very low, perceived design technique knowledge did not influence selection accuracy. Thus, in this experiment, as the experiment participants were also novices, the internal problem representation was not considered as a predictor when developing hypotheses.

[3] As in part 2 of cycle 2 (Chapter 5.2), intuitive decision-making style did not influence selection accuracy, in this experiment, intuitive decision-making style was not considered as a predictor when developing hypothesis.

on analyzing the support of two UIs on the design novices' selection of design techniques, the two UIs are representations of a decision aid. As the novices have no previous knowledge of the UIs for supporting the selection of design techniques, one can assume that the experiences of using the experimental tool will be accumulated with the increase of the repeated use of it. The more times people use the tool, the more familiar people should become with the functions and information included in the UIs. And the more times people use the tool; the less time they should spend on the task. Thus, the more tasks are conducted by using the experimental tool; the less time will be used on each individual task. Thus, both treatments are assumed to reduce the time required for each task with an increase of the overall number of tasks performed:

> ***H1a:*** *Using a taxonomy UI leads to a decrease of time used in each task with the increase of the number of tasks.*

> ***H1b:*** *Using a natural language UI leads to a decrease of time used in each task with the increase of the number of tasks.*

Influence of UIs on Selection Accuracy and Perceived Task Performance

In this research, the selection criteria (i.e., the dimensions and categories of design techniques) in the taxonomy UI and the natural language UI are the same, only the visual presentation is different. As different visualizations of the same information have an impact on the decision-making result (Lurie and Mason 2007), the selection accuracy and perceived task performance by using the different UIs may be different. When making choices, people tend to compare different alternatives in the decision-making process before making the final decision (Karimi et al. 2015). Thus, selection behavior can be matched by providing decision-makers with the possibilities of comparing alternative design techniques. Visualizing all the categories of design techniques in the taxonomy UI with structured dimensions and characteristics with all the available design techniques correspond to the decision-making behavior. In contrast, the natural language UI reduces the possibility of comparing all the alternatives. Although the natural language UI explains the categories of the taxonomy in an easier way, the visualization does not match the selection behavior. According to cognitive fit theory, when there is a fit between the task and the problem representation, the task performance will increase (Vessey and Galletta 1991). Hence, when using a taxonomy UI, the selection accuracy and task performance are expected to be higher than using a natural language UI.

> ***H2:*** *Using a taxonomy UI leads to higher selection accuracy than a natural language UI.*

> ***H3:*** *Using a taxonomy UI leads to higher perceived task performance than a natural language UI.*

Influence of UIs on Perceived Complexity and Ease of Use moderated by Rational Decision-making Style

In the decision-making process, when using the UIs to find the appropriate design techniques, different perceptions of complexity and ease of use are derived. Perceived complexity as well as ease of use are subjective evaluation for the UIs (Davis 1989; Thompson et al. 1994). Thus, differences in decision-making style can influence the effect of different UIs on the individual subjective evaluation (Simon and Usunier 2007). Commonly, differences are observed between low and high rational decision-makers (Hamilton et al. 2016). In a decision-making task, as high rational decision-makers are analytic and make decisions based on evidential factors, they prefer to collect all available information and compare alternatives (Cacioppo et al. 1984; Hamilton et al. 2016). Thus, when all the information and selection possibilities are presented to high rational decision-makers in a UI of a decision aid, the perception of complexity will decrease, and ease of use will increase. However, too much information may get low rational decision-makers overwhelmed, which in turn leads to an increase of perceived complexity and a decrease of perceived ease of use.

> ***H4a:*** *Rational decision-making style has a moderating effect on the influence between using the taxonomy and natural language UIs and perceived complexity.*

> ***H4b:*** *Rational decision-making style has a moderating effect on the influence between using the taxonomy and natural language UIs and perceived ease of use.*

Influence of Perceived Complexity and Ease of Use on Task Performance

Unlike the problem complexity (Mennecke et al. 2006), perceived complexity is the perception of the complexity resulting from information overload, time pressure, and other environmental factors (Kirk et al. 2015; Ordóñez and Benson 1997). In this study, the perceived complexity is expected to result from using the experimental tool with the two treatments. Research in website design shows that visual complexity has a negative effect on task performance (Tuch et al. 2009). Thus, in the decision-making task, perceived complexity is predicted to have a negative effect on task performance. Similar to perceived complexity, perceived ease of use is the subjective evaluation of the tool used when conducting the task (Davis 1989). In contrast to the effect of perceived complexity, an increase of perceived ease of use can enhance task performance (Wixom and Todd 2005). Hence, when using different UIs for selecting design techniques, perceived ease of use is predicted to positively influence the perceived task performance.

> ***H5a:*** *When using the taxonomy and natural language UIs to select design techniques, perceived complexity has a negative effect on perceived task performance*

> ***H5b:*** *When using the taxonomy and natural language UIs to select design techniques, perceived ease of use has a positive effect on perceived task performance*

The proposed hypotheses are embedded in a research model using the already introduced IPO framework. The two different UIs are inputs; the perceived complexity and perceived ease of use as well as the rational decision-making style belong to process; Time used, selection accuracy, and perceived task performance of the output. Figure 6-10 depicts the research model.

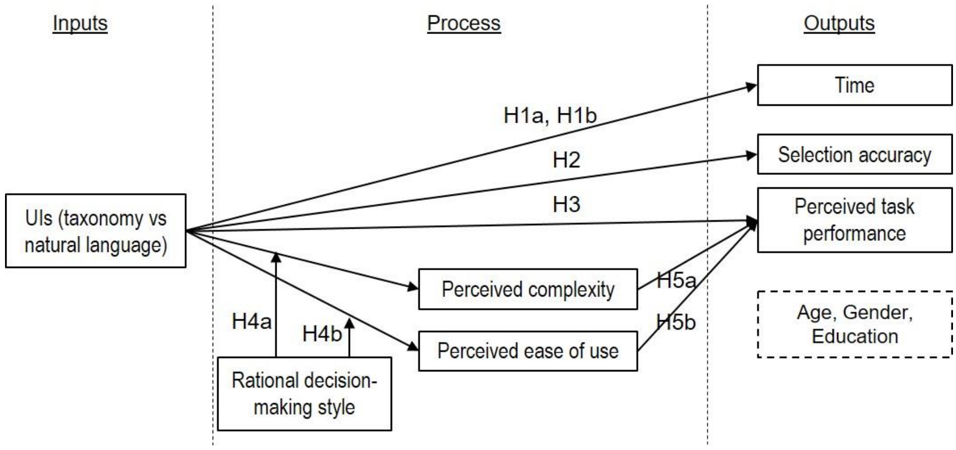


Figure 6-10. Research model for the hypotheses of the effects of different UIs

6.2.3 Lab Experiment

In order to test the proposed hypotheses, a between-subject (taxonomy vs. natural language dialog) lab experiment was conducted. In the following, further details on the experiment design are described.

Experiment Design

Experiment Setting

An experimental tool was developed, which included a taxonomy (Figure 6-11) and a natural language UI (Figure 6-12) to support the experiment participants to complete the experimental task. The classified design techniques and the classification were adopted by Liu et al. (2016)

(part 1 in cycle 1, Chapter 4.1). The classification contents were identical; the only difference is the visualization of the categories. The taxonomy used abstract terms for the dimensions and characteristics, while the natural language UI explained the abstract terms in natural language. When using the taxonomy UI, experiment participants clicked on the characteristics to filter design techniques. Whereas, when using the natural language UI, the experiment participants answered several questions that explain those dimensions and characteristics to get the filtered design techniques based on their answers. Each experiment participant used only one of the UIs for completing the task. In addition, as each of the experiment participants conducted the experiment in a soundproofed booth of the KD2Lab, there was no disruption from others. The task used for selecting design techniques was the same in experiment of part 2 of cycle 2 (Chapter 5.2). Detail description can be found in Appendix G.

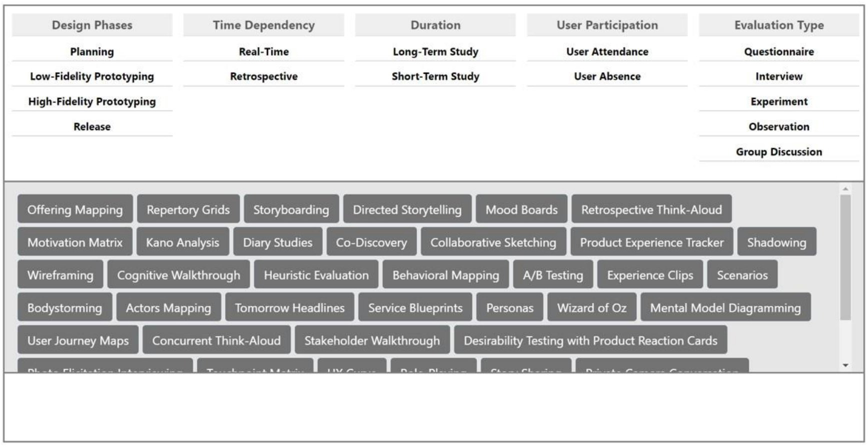


Figure 6-11. Using taxonomy UI to select design techniques

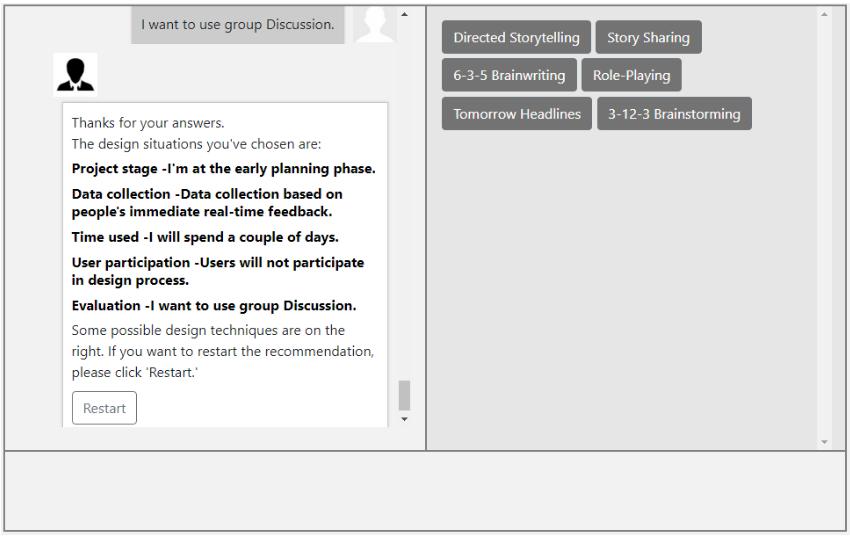


Figure 6-12. Using natural language UI to select design techniques

Experiment Procedure

The experiment process included three stages (Figure 6-13). There were two groups, the taxonomy UI group, and the natural language UI group.

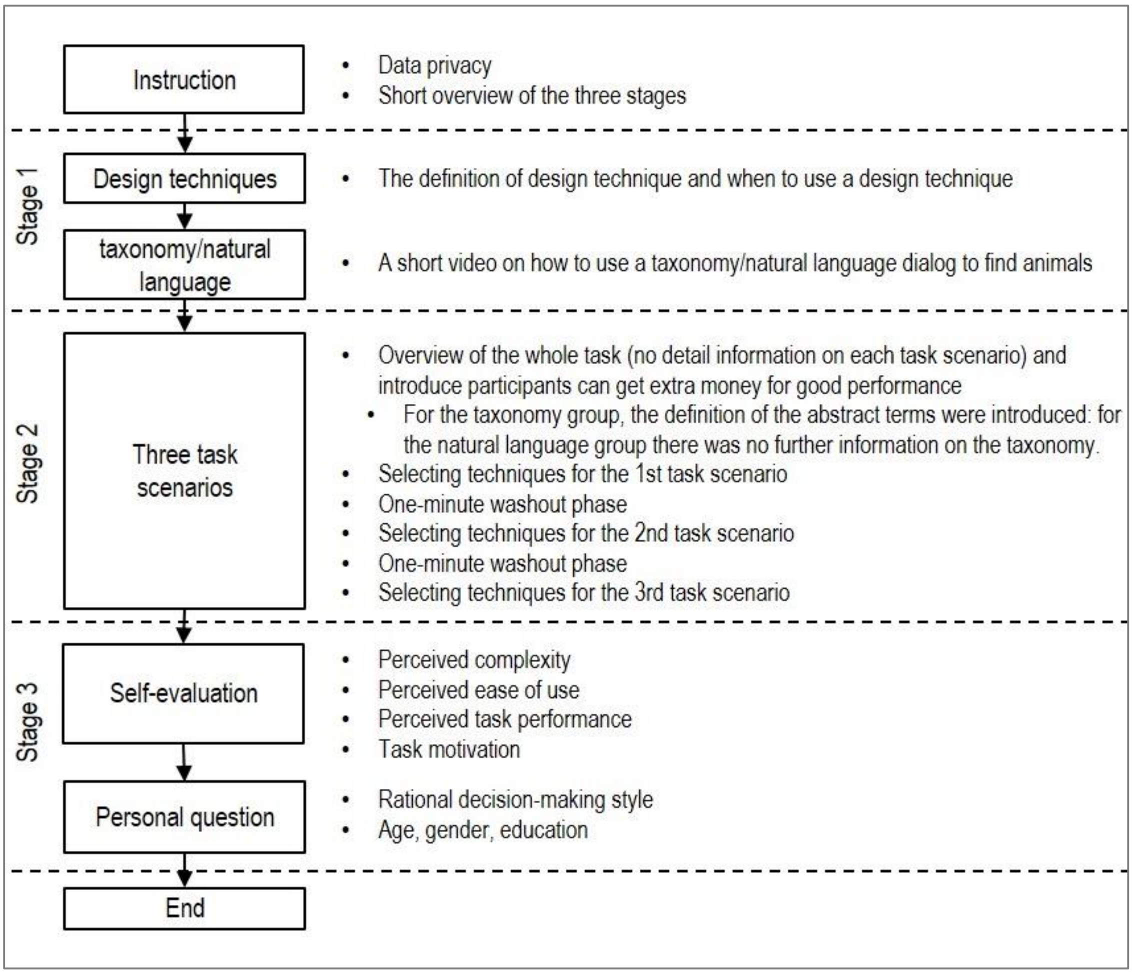


Figure 6-13. Experiment procedure for evaluating the effect of different UIs

Before introducing the experiment tool, an instruction of data privacy and an overview of the three stages were at first presented to the experiment participants. *Stage 1* was a short training part, which at first introduced the definition of design techniques and when a design technique can be used. Subsequently, a one-minute video that introduced the experiment was presented (Appendix M). the taxonomy-based UI tool and natural language-based UI tool were assigned to the corresponding groups separately. In order not to bias the experiment participants, no information of design techniques and design tasks were mentioned in the introduction. Instead, an example of finding animals using a classification was used. At the end of each introduction, there was a question on whether the participant understood the introduction. If a participant chose no, the data point was removed in the data analysis. *Stage 2* presented the experiment task and collected the selected design techniques for the task from the participants. For the taxonomy UI group, the definition of the dimensions and characteristics in the taxonomy of design techniques was provided. The time used for reading the information on the dimensions and characteristics of the taxonomy was recorded. Whereas, as all the abstract terms were explained in the natural language tool, there was no extra information on the taxonomy of design techniques to the natural language group. In order to test whether time used can change with the increase of the number of tasks, three tasks were used in the experiment (Appendix G). Between two tasks, there was a one-minute washout phase. The time used for each task was recorded. The three tasks and the two washout phases were totally the same for the taxonomy group and the natural language group. But the two groups used different tools when completing the task. The taxonomy UI group only used the taxonomy for the three task scenarios, while the natural language group only used the natural language UI for the three task scenarios. For each task scenario, a participant was allowed to select maximum five design techniques. In *stage 3*, self-evaluation on perceived complexity (Thompson et al. 1994), perceived ease of use (Wixom and Todd 2005), and perceived task performance (Kim et al. 2009) were measured. In addition, personal questions on rational-decision-making style (Hamilton et al. 2016) and demographic questions, such as age, gender, education, were asked.

During the experiment, screen recordings of all the experiment participants during the experiment were made, which were used to remove the invalid data such as participants who quickly clicked through the questions in the experiment.

Variables and Measurements

As this experiment focused on the evaluation of the different effect of UIs on selection performance, the treatments included taxonomy and a natural language dialog to support the selection of design techniques. Table 6-5 summarizes the definitions of UIs of taxonomy and natural language dialog when applying them in the selection of design techniques used in this research.

Table 6-5. Definitions of the treatments used in the experiment

Decision Aids	Definition	Source
Taxonomy UI	A taxonomy UI is visualized with hierarchical structured and consists of abstract dimensions and characteristics.	(Nickerson et al. 2013)

Natural language UI	A natural language UI is visualized with pre-defined questions in natural language which explains the dimensions and characteristics in the taxonomy.	(Arora et al. 2013; Chen et al. 2017; McTear 2002)

Based on the developed hypotheses, the rational decision-making style refers to the degree to which people tend to collect all information and comparing all alternatives when making decisions, which is an independent variable and measured by using the items developed by Hamilton et al. (2016). The other variables, time used, selection accuracy, perceived complexity, perceived ease of use, and perceived task performance were used as dependent variables which were predicted to be influenced when using the two different UIs. Time used was hypothesized to decrease with an increasing number of completed tasks. Thus, time used needed to be measured for each of the three tasks individually. For the taxonomy UI group, as there was additional information on the explanation of the dimensions and characteristics of the taxonomy, the time used for reading such information was added to the time used in the task. As the information of the taxonomy was only presented once before the first task, the time used for reading the information was added to the time used for the first task. For the natural language UI group, time used was the time spent on each task. Except for time used, selection accuracy did not need to be measured with a dedicated construct. As for each task, a participant was asked to select at most 5 design techniques for the task scenario, the highest accuracy was 15 when all the selections were correct. The score of selection accuracy ranges from 0 to 15. For example, if an experiment participant provides 2 correct selections for the first scenario, 3 correct selections for the second one, and 5 correct selections for the last one, the resulting selection accuracy will be 9. Hence, selection accuracy was measured by the number of correct selections in the three tasks. The other dependent variables were measured through dedicated measurements. Perceived complexity describes the extent to which people think a system is complex and difficult to use (Thompson et al. 1994). The measurement of perceived complexity was adopted from Thompson et al. (1994). Perceived ease of use describes the degree to which people think a system is easy to use (Davis 1989). The measurement for perceived ease of use was adopted from Wixom and Todd (2005). Finally, perceived task performance is defined as the degree to which people think a system help with the task performance (Churchill and Surprenant 1982; Kim et al. 2009). The measurement for perceived task performance was adopted from Kim et al. (2009). All items in the measurements were measured using a 7-point Likert scale (Burton-Jones and Straub 2006) (Appendix N). A summary of the variables and definitions is presented in Table 6-6.

Table 6-6. Definitions of variables used in the experiment

Variables	Definition	Source
Time used	Time spent on each selection task in the experiment	-
Accuracy	The number of correct selections of design techniques	-
Perceived complexity	The participants' perception of the degree to which using the taxonomy, or the natural language dialog would be complex and difficult	(Thompson et al. 1994)
Perceived ease of use	The participants' perception of the degree to which using the taxonomy, or the natural language dialog would be free of effort	(Davis 1989)
Perceived task performance	The participants' perception of how the taxonomy or the natural language dialog helped them in their task performance	(Churchill and Surprenant 1982; Kim et al. 2009)
Rational decision-making style	The degree to which a participant makes decisions based on a thorough search of all information and comparison of alternatives	(Hamilton et al. 2016)

Participants and Incentives

148 participants[1] were recruited from KD2Lab Hroot[2] panel to participate in the experiment. After the experiment, eight data points were removed because of incomplete answers, misunderstanding of the task explanation, or quickly clicking through the questions used as measurements in the analysis. As a result, 140 valid data points (70 for natural language and 70 for taxonomy) were used for the experiment analysis. 60 were female, 78 were male, 2 answered other. The average age was 23.39 years (SD = 2.82). 98% of the participants were students. As the purpose of this experiment was to understand whether different UIs could influence the novices' selection of design techniques, students were considered as an appropriate group for the experiment (Druckman and Kam 2012). Each participant was assigned to one of the two UIs randomly. 10 sessions were conducted in three days. Each participant received at least 7 Euros as compensation for the participation. In order to incentivize the participants to take the selection seriously, participants could earn an additional 3 Euros based on their performance in the experiment. For each task, there were 7 correct selections. As a maximum of 5 selections were allowed, if a

[1] The sample size was calculated with the tool "G*Power 3.1.9.2" with an a priori test (Faul et al. 2007). When considering the statistical power of 0.80 for a medium effect size (f=0.25) and a significance level of 0.05 (Cohen 1988), the minimum sample size should be 128. Thus, we recruited 148 people to participate the experiment to make sure there were enough subjects in the experiment. After the experiment, eight data points were removed because of incomplete, mistake operation, or mistaken understanding of the experiment tasks.

[2] https://iism-kd2-hroot.iism.kit.edu/?locale=en

participant would provide at least 4 correct selections, they received 1 Euro extra. Over the course of three tasks, an extra of 3 Euros could be earned.

6.2.4 Data Analysis and Results

The data analysis was performed in two steps: 1) measurement validity and 2) hypotheses test. In the measurement analysis, an EFA was conducted to test the factor loading and Cronbach alphas of measurements. The correlations between the variables were also analyzed. In the hypotheses test, the statistical analysis and results were presented based on the sequence of hypotheses development.

Measurement Validity

An EFA using principal component analysis and varimax rotation was performed on perceived task performance, rational decision-making style, perceived complexity, and perceived ease of use (cf., Yi et al. 2015). The items loaded highly on their intended factor and lowly on the other factors (Appendix O). The Cronbach alphas of the measure of perceived task performance, rational decision-making style, perceived complexity, and perceived ease of use were 0.9, 0.85, 0.88, and 0.81 respectively, indicating adequate reliability of the measurement scales. The correlations between the four variables measured by measurements and selection accuracy are reported in Table 6-7.

Table 6-7. Correlations between variables

Variables	Mdn	M	SD	Correlations				
				1	2	3	4	5
1. Selection Accuracy	12	12	3.77	1				
2. Perceived Ease of Use	6	5.90	0.88	0.30***	1			
3. Perceived Task Performance	5.75	5.54	1.05	0.37***	0.59***	1		
4. Perceived Complexity	2	2.22	1.07	-0.42***	-0.58***	-0.47***	1	
5. Rational Decision-making	5.6	5.44	0.87	0.16+	0.31***	0.31***	-0.22+	1

*Note: +P<0.1; *P<0.05; **P<0.01; ***P<0.001*

Hypotheses Test

Before starting hypotheses test, a Shapiro-Wilk test was conducted to check the normalization of the variables. The result presented that rational decision-making *($p < 0.05$)*, selection accuracy *($p < 0.05$)*, perceived ease of use *($p < 0.05$)*, perceived complexity *($p < 0.05$)*, perceived task performance *($p < 0.05$)*, and the time used in three tasks *($p < 0.05$)* were not normally distributed. Thus, Friedman's ANOVA was used to test whether there were differences between the time used in three tasks (H1a and H1b); Wilcoxon Rank-sum test was used to whether different UIs lead to different selection accuracy and perceived task performance (H2 and H3). As H4a and H4b were an assumption of the moderating effect of rational decision-making style, linear hierarchical regression analysis was used. For H5a and H5b, multiple linear regression was used. In the following, detail descriptions on analyses and results are presented.

Influence of UIs on Time used for different Tasks

A summary of the test for H1a and H1b is presented in Table 6-8. Both hypotheses were analyzed by using a within-subject repeated test.

Table 6-8. A summary of used tests and the results of H1a and H1b

#	Hypothesis	Test	Result
H1a:	Using a taxonomy UI leads to a decrease of time used in each task with the increase of the number of tasks	Friedman's ANOVA	Supported
H1b:	Using a natural language UI leads to a decrease of time used in each task with the increase of the number of tasks	Friedman's ANOVA	Not supported

For testing H1a, Friedman's ANOVA was applied. There is a significant direct effect of using the taxonomy-based UI on the time used in three tasks *[x^2 (2) = 91.91, p < 0.000]*. In order to determine the time used in which tasks differ from the others, a post hoc test with Bonferroni correction was applied, which showed that there were significant differences between Task 01 and Task 02 *(difference = 79)*, and between Task 02 and Task 03 *(difference = 31)*, as well as between Task 01 and Task 03, *(difference = 110)*. In all cases, the critical difference ($\alpha = 0.05$ corrected for the number of the tests) was 28.33. The result shows that the more times people use the taxonomy-based UI, the less time would be required. Hence H1a was supported.

The same tests were applied to test H1b. Similar to the time used in three tasks by using a taxonomy-based UI, a significant direct effect of using a natural language dialog-based UI on the time used in three tasks was presented *[x^2 (2) = 51.51, p <0 .000]*. But the post hoc test showed that the comparisons of time used in three tasks were not all significantly different. There was a significant difference between time used in Task 01 and Task 02 *(difference = 75)*, and between Task 01 and Task 03 *(difference = 72)*. However, the time used for Task 02 was not significantly different from Task 03 *(difference = 3)*. In all cases, the critical difference ($\alpha = 0.05$ corrected for the number of the tests) was 28.33. The result reflected that there was a decrease in time used between the first two tasks, but no difference between the second and third task, which meant the time used would not decrease with the increase of the number of tasks. Hence, H1b was not supported.

Further comparison between the time used in each task across using the two different UIs was conducted. A Wilcoxon Rank-sum test was used to analyze the differences of time used. The result indicated that for Task 01, time used by using a natural language UI *(Mdn = 423.00)* was significantly lower than using a taxonomy UI *(Mdn = 490.94), p =0 .029, r = -0.19*. But there was no significant difference between time used for Task 02 when using natural language dialog UI *(Mdn = 284.48)* and taxonomy UI *(Mdn = 271.80), p = 0.502, r = -0.057* Whereas, in Task 03, the time used by using a taxonomy UI *(Mdn = 189.98)* was significantly lower than a natural language UI *(Mdn = 256.63), p = 0.0017, r = -0.27.*

The boxplot of the time used in the three tasks by using taxonomy and natural language dialog UI is presented in Figure 6-14.

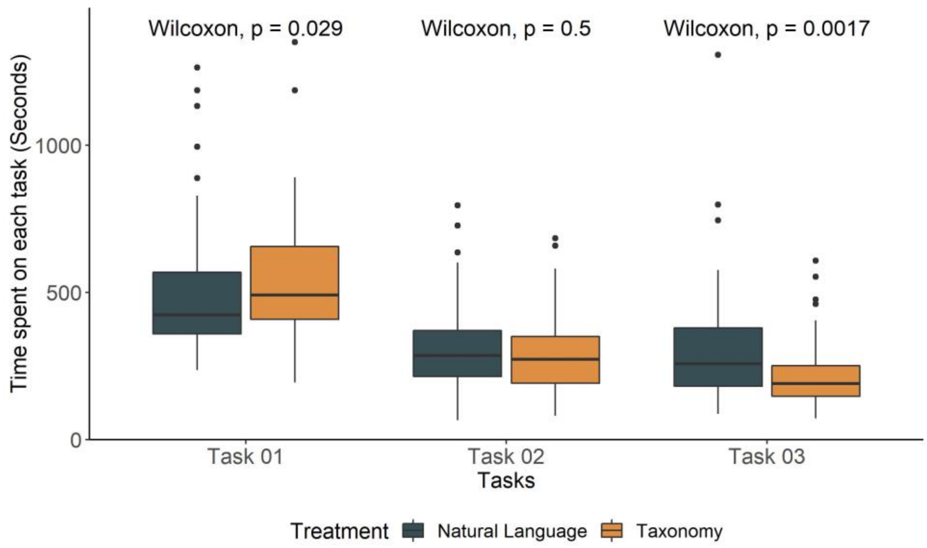


Figure 6-14. Boxplot for the time used for the three tasks across different UIs

Influence of UIs on Selection Accuracy and Perceived Task Performance

A summary of the test for H2 and H3 is presented in Table 6-9.

Table 6-9. A summary of used tests and the results of H2 and H3

#	Hypothesis	Test	Result
H2:	Using a taxonomy UI leads to higher selection accuracy than using a natural language UI	Wilcoxon Rank-sum test	Supported
H3:	Using a taxonomy UI leads to higher perceived task performance than using a natural language UI	Wilcoxon Rank-sum test	Supported

The Wilcoxon Rank-sum test result indicated that the selection accuracy by using a natural language UI *(Mdn = 12)* was significantly lower than using a taxonomy UI *(Mdn = 12), p = 0.032, r = -0.18.* Similarly, the perceived task performance by using a natural language UI *(Mdn = 5.25)* was significantly lower than using a taxonomy UI *(Mdn = 6), p < 0.001, r = -0.29.* Hence, H3 and H4 were supported. The boxplots of the comparison of selection accuracy and perceived task performance by using taxonomy and natural language UI are presented in Figure 6-15.

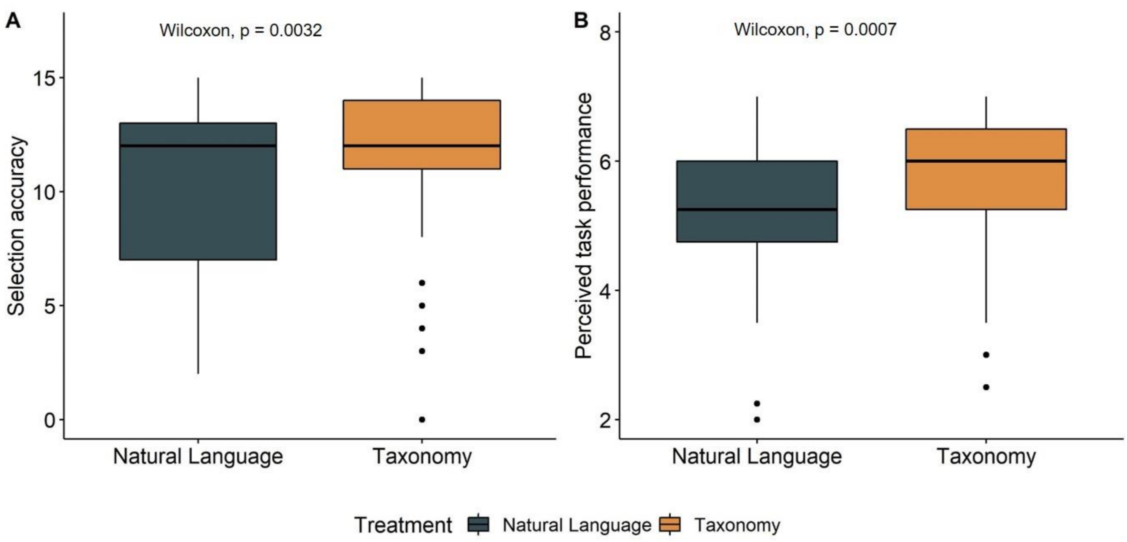


Figure 6-15. Comparison of selection accuracy (A) and perceived task performance (B)

Moderating Effect of Rational Decision-making Style

A summary of the test for H4a and H4b is presented in Table 6-10.

Table 6-10. A summary of used tests and the results of H4a and H4b

#	Hypothesis	Test	Result
H4a:	Rational decision-making style has a moderating effect on the relationship between using the two UIs and perceived complexity	Hierarchical linear regression	Supported
H4b:	Rational decision-making style has a moderating effect on the relationship between using the two UIs and perceived ease of use	Hierarchical linear regression	Supported

Before testing the moderating effect of rational decision-making style, a Wilcoxon Rank-sum test was used to test whether there was no difference in the distribution of rational decision-making style across the taxonomy UI group and the natural language UI group. The result showed that there was no difference between the taxonomy group *(Mdn = 5.5)* and natural language group *(Mdn = 5.6)*, $p = 0.335$, $r = -0.081$. Thus, the rational decision-making style can be used for further regression test across the two groups. A hierarchical linear regression test (Baron and Kenny 1986) was used to analyze the moderating effect of rational decision-making style. The result of the regression on perceived complexity (H4a) and perceived ease of use (H4b) is summarized in Table 6-11.

For testing H4a, after testing all control variables (Step 1), a regression model without the moderating effect of rational decision-making style (Step 2) was built, which showed that rational decision-making style significantly regressed onto the perceived complexity $(p = 0.007)$, but the two different UIs did not significantly regress on the perceived complexity $(p = 0.89)$. In addition, the regression model in Step 2 was not significant $(P = 0.12)$. In Step 3, a regression model with the moderating effect of rational decision-making style was built, which presented the UIs, rational decision-making style, and their interaction effect all significantly regressed onto the perceived complexity, with $p < 0.001$, $p = 0.010$, $p< 0.001$ separately. With adding the moderating effect of rational decision-making style, the R^2 of the model in Step 2 and 3 increased from *0.062* to *0.144*. An ANOVA test of the models in Step 2 and 3 showed that the two models were significantly different $(p < 0.001)$. Hence H4a was supported.

For testing H4b, in Step 5, a regression model without the rational decision-making style as a moderator was built. It demonstrated that rational decision-making style significantly regressed on the perceived ease of use *($p < 0.001$)* but the two different UIs did not significantly regress onto the perceived ease of use *($p = 0.80$)*. In Step 6, a regression model with the moderating effect of rational decision-making style was built, which presented the UIs, rational decision-making style, and their interaction effect all significantly regressed onto the perceived ease of use, with $p < 0.001$, $p = 0.014$, $p < 0.001$ separately. After adding rational decision-making style as a moderator, the R^2 of the model in Step 5 and 6 increased from *0.097* to *0.183*. An ANOVA test of the regression model without and with the moderating effect of rational decision-making style presented that the two models were significantly different *($p < 0.001$)*. Hence, H4b was supported. The robust test result of regression models of Step 3 (H4a) and Step 6 (H4b) is in Appendix P.

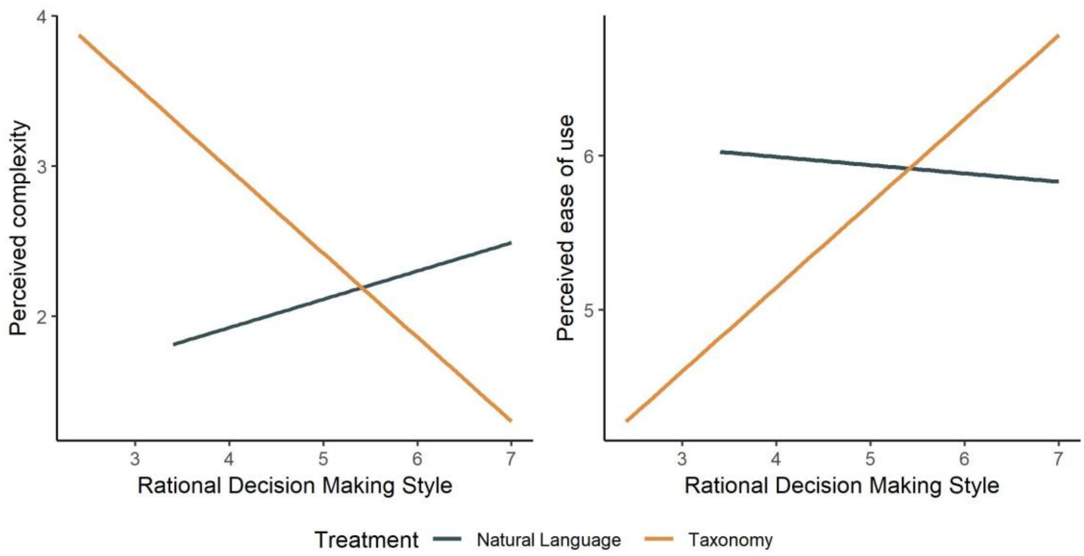
Table 6-11. Hierarchical regression test of H4a and H4b (n = 140)

	Dependent variable:					
	Perceived complexity			***Perceived ease of use***		
Model	Step 1: Control variables	Step2: without moderator	Step 3: with moderator (H4a)	Step 4: Control variables	Step 5: without moderator	Step 6: with moderator (H4b)
Age	0.013 (0.038)	0.031 (0.038)	0.023 (0.037)	0.028 (0.031)	0.009 (0.031)	0.016 (0.029)
Gender	0.156 (0.174)	0.133 (0.171)	0.041 (0.166)	-0.031 (0.143)	-0.006 (0.138)	0.072 (0.133)
Edu	0.037 (0.175)	0.020 (0.173)	0.073 (0.166)	-0.041 (0.144)	-0.023 (0.139)	-0.068 (0.133)
UI		-0.025 (0.181)	**4.011*** (1.144)**		0.037 (0.146)	**-3.365*** (0.918)**
RDS		**-0.287** (0.105)**	**0.908* (0.349)**		**0.310*** (0.085)**	**-0.697* (0.281)**
UI X RDS			**-0.739*** (0.207)**			**0.623*** (0.166)**
Constant	1.620* (0.799)	2.866** (0.988)	-3.442+ (2.004)	5.351*** (0.658)	3.983*** (0.797)	9.300*** (1.610)
R^2	0.009	**0.062**	**0.144**	0.006	**0.097**	**0.183**
ΔR2	-	**0.053**	**0.135**	-	**0.091**	**0.177**
F Statistic	0.425 (df = 3; 136)	1.767 (df = 5; 134)	3.726** (df = 6; 133)	0.280 (df = 3; 136)	2.876* (df = 5; 134)	4.972*** (df = 6; 133)

*Notes: 1) the table gives coefficients (standardized errors); 2) +P<0.1; *P<0.05; **P<0.01; ***P<0.001; 3) RDS: rational decision-making style; 4) ΔR^2 scores compare the model against the controls only model.*

Figure 6-16 visualizes the regression lines indicating that the relationship between the use of different UIs and the perceived complexity as well as perceived ease of use were moderated by rational decision-making style. The regression lines in the left plot shows that when using a natural language UI, the perceived complexity increases with the increase of the level of rational decision-making style. Whereas, when using a taxonomy UI, the perceived complexity reduces with the increase of the level of rational decision-making style. The regression line in the right plot shows that the higher the rational, the higher perception of the ease of use when using a taxonomy UI.

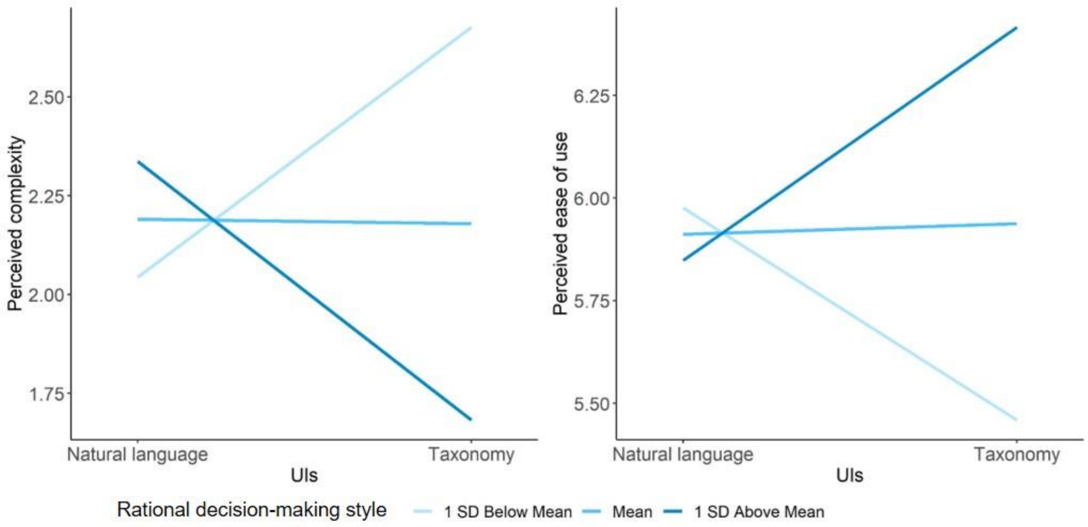
However, there was a slight drop of perceived ease of use with the increase of the level of rational decision-making style when using a natural language UI.



Notes: moderating effect of rational decision-making style between the relation of two UIs and perceived complexity (left) and perceived ease of use (right)

Figure 6-16. Linear regression of the moderating effect of rational decision-making style

The simple slopes for perceived complexity and ease of use were divided at three levels of rational decision-making style (1SD above the mean, mean, and 1SD below the mean) (Figure 6-17) (Aguinis and Gottfredson 2010; Artz et al. 2012). For low rational decision-makers (1 SD below the mean), when compared with natural language UI, taxonomy had a positive effect *(beta = 1.60, t = 0.44, p < 0.001)* on perceived complexity, while the interaction effect of taxonomy and rational decision-making style had a negative effect *(beta = -1.64, t = 0.44, p < 0.001)* (left of Figure 6-17). In addition, the effect of taxonomy on perceived ease of use for low rational decision-makers (right of Figure 6-17) was negative *(beta = -1.55, t = 0.43, p < 0.001)*, while the effect of the interaction effect of taxonomy and rational decision-making style was positive *(beta = 1.60, t = 0.43, p < 0.001)*. For high rational decision-makers (1 SD above the mean), when compared with natural language UI, taxonomy *(beta = 2.21, t = 0.61, p < 0.001)* had a positive effect on perceived complexity (left of Figure 6-17), while the interaction effect of rational decision-making style and taxonomy was negative *(beta = -2.24, t = 0.61, p < 0.001)*. Whereas, the effect of taxonomy on perceive ease of use for high rational decision-makers (right of Figure 6-7) was negative *(beta = -2.15, t = 0.59, p < 0.001)*, while the interaction effect was positive *(beta = 2.18, t = 0.59, p < 0.001)*.



Notes: moderating effect of rational decision-making style between the relationships of UIs and perceived complexity (left) and ease of use (right)

Figure 6-17. SD analysis of the moderating effect of rational decision-making style

In order to have a deeper look at the differences between the perception of complexity and ease of use with low- and high-level of rational decision-making style, two data subsets were created based on the median of rational decision-making style *(Mdn = 5.6)*. The participants with a level of rational decision-making style lower than 5.6 were assigned to the subset of the low rational group (31 natural language and 35 taxonomy). The others were assigned to the high rational group (39 natural language and 35 taxonomy). The differences between perceived complexity and ease of use were analyzed in each subset separately. Figure 6-18 shows the differences in the perception of complexity and ease of use by using taxonomy UI and natural language UI across low- and high-level of rational decision-making styles.

A Wilcoxon Rank-sum test was used to test the differences in each subset. For perceived complexity, the result presented that with a low-level of rational decision-making style (A in Figure 6-18), using natural language UI *(Mdn = 1.75)* led to lower perceived complexity than a taxonomy UI *(Mdn = 2)* with a trend toward significance, $p = 0.058$, $r = -0.23$. Whereas, with a high-level of rational decision-making style (B in Figure 6-18), the perceived complexity was significantly lower when using taxonomy UI *(Mdn = 1.75)* than natural language UI *(Mdn = 2)*, $p = 0.003$, $r = -0.34$. For perceived ease of use, with a low-level of rational decision-making style (C in Figure 6-18), using taxonomy UI *(Mdn = 6)* led to lower perceived complexity than natural language UI *(Mdn = 6)* at the margin of statistical significance, $p = 0.063$, $r = -0.23$.

While, with a high-level of rational decision-making style (D in Figure 6-18), the perceived ease of use by using natural language UI *(Mdn = 6)* was significantly lower than using taxonomy UI *(Mdn = 6.33), p = 0.008, r = -0.31*.

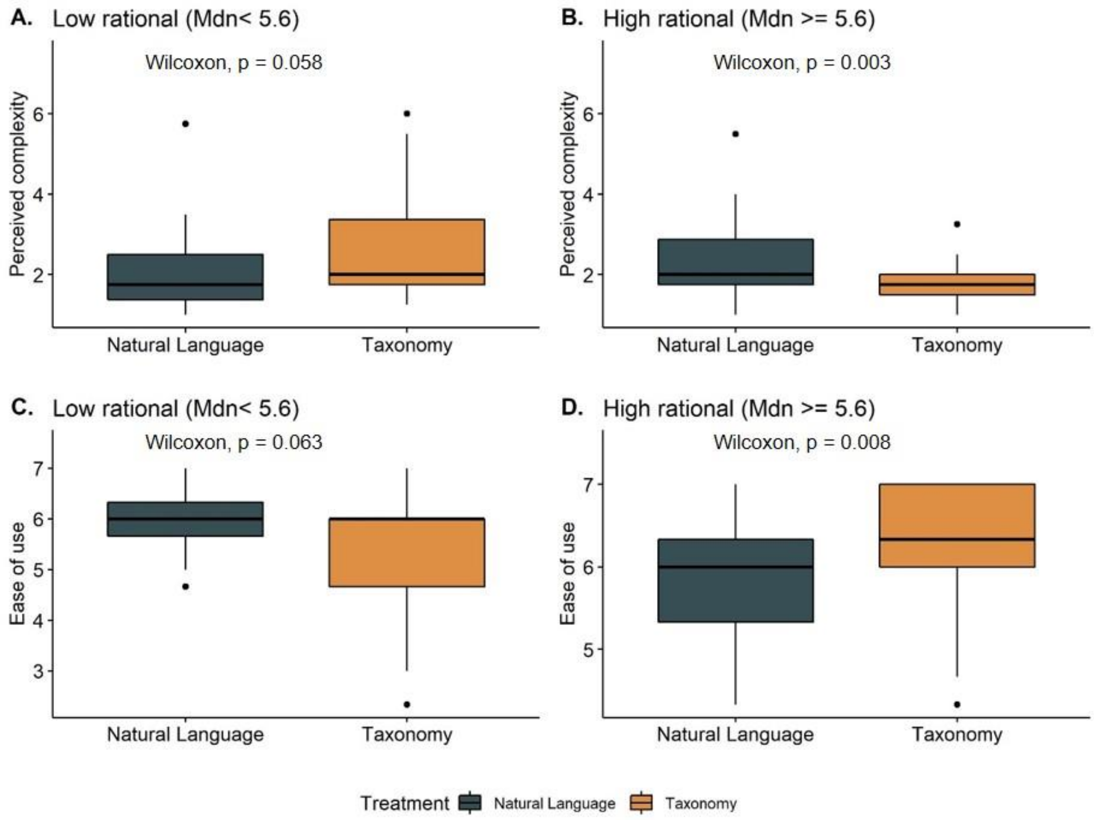


Figure 6-18. Boxplots for the comparison of perceived complexity and ease of use

Influence of Perceived Complexity and Ease of Use on Task Performance

A summary of the tests for H5a and H5b is presented in Table 6-12.

Table 6-12. A summary of used test and the result of H5a and H5b

#	Hypothesis	Test	Result
H5a:	When using the taxonomy and natural language UIs to select design techniques, perceived complexity has a negative effect on perceived task performance	Hierarchical regression	Supported
H5b:	When using the taxonomy and natural language UIs to select design techniques, perceived ease of use has a positive effect on perceived task performance	Hierarchical regression	Supported

Multiple linear regression was conducted to test H5a and H5b (Table 6-13). In Step 2 (H5a), a linear regression was calculated to predict perceived task performance based on perceived complexity. The result showed a negative significant effect of perceived complexity with $P < 0.001$, $R^2 = 0.230$. In Step 3 (H5b), perceived ease of use had a positive effect on perceived task performance, with $P < 0.001$, $R^2 = 0.351$. Hence, H5a and H5b were supported. In Step 4, further regression analysis showed that effect of both perceived complexity and ease of use on perceived task performance was significant, with $p < 0.023$, $p < 0.001$ separately and $R^2 = 0.375$. The robust test of H5a and H5b is in Appendix P.

Table 6-13. Hierarchical regression results for testing H5a and H5b (n = 140)

	Dependent variable: performance			
Model	Step 1: Control variables	Step 2: H5a	Step 3: H5b	Step 4: H5a & H5b
Age	0.041 (0.037)	0.047 (0.033)	0.022 (0.030)	0.028 (0.030)
Gender	-0.036 (0.171)	0.037 (0.152)	-0.014 (0.139)	0.011 (0.137)
Edu	-0.087 (0.173)	-0.070 (0.153)	-0.058 (0.140)	-0.057 (0.138)
PC		**-0.465*** (0.075)**		**-0.192* (0.083)**
PEU			**0.701*** (0.083)**	**0.565*** (0.101)**
Constant	4.756*** (0.786)	5.508*** (0.706)	1.002 (0.779)	2.044* (0.890)
R^2	0.009	**0.230**	**0.351**	**0.375**
ΔR2	**-**	**0.221**	**0.342**	**0.366**
F Statistic	0.409 (df = 3; 136)	10.085*** (df = 4; 135)	18.226*** (df = 4; 135)	16.101*** (df = 5; 134)

*Notes: 1) the table gives coefficients (standardized errors); 2) +P<0.1; *P<0.05; **P<0.01; ***P<0.001; 3) PC: perceived complexity; PEU: perceived ease of use; 4) ΔR2 scores compare the model against the controls only model.*

6.2.5 Discussion of Results

This section discusses the results from the experimental study introduced above. Specifically, the differences of using a natural language or a taxonomy UI for short-term or long-term use based on the differences of time used across tasks is reflected. Additionally, the role of rational decision-making is discussed when using a natural language or a taxonomy UI based on different perception of complexity and ease of use.

Applying Natural Language UI for Short-term Use but Taxonomy UI for Long-term Use

The experiment result shows that when using a natural language UI, the time used for the first task is significantly lower than the time used when using a taxonomy UI. The reason is that people need to read the information such as the definition of the dimensions and characteristics in the taxonomy in order to understand the taxonomy before using it for the selection task. Whereas in the natural language UI, the abstract terms were explained which allowed for the omission of the terminology explanation step prior to the selection. Such a feature of the natural language UI led to the lower time used for the first task. However, in the second task, there already observed no significant difference. In the third task, time used with the natural language UI became higher than using the taxonomy UI. This phenomenon may be explained through the increasing information understanding when using the taxonomy UI as users repeatedly use the tool. Although users also became more familiar with the natural language UI, the same questions were asked and needed to be answered for the three tasks. Compared with the taxonomy UI, the natural language UI reduces the initial information presented to people, which can benefit the selection process when needed for short-term use. On the contrary, when using a taxonomy UI, the speed in using the tool increases with growing understanding and experience. As such, the taxonomy UI is beneficial for repeated selection processes, e.g. in long-term use cases.

The Role of Rational Decision-making Style when Using Natural Language UI or Taxonomy UI

The hypotheses test shows that perceived complexity and perceived ease of use are influenced by different UIs only under consideration of the moderating effect of rational decision-making style. The plotted regression lines (Figure 6-16 and 6-17) and the deeper analysis by creating subsets of participants with low and high-level of rational decision-making style (Figure 6-18) indicates that people with different rational decision-making styles have different perceptions of complexity and ease of use when using natural language UI and taxonomy UI. Low rational decision-makers had a higher perception of the complexity and lower perception of ease of use when using the taxonomy UI. Cognitive fit theory can be used to explain the result (Shaft and Vessey 2006; Vessey and Galletta 1991). Information processing process of high rational decision-makers involves collecting all available information and comparing all alternatives before making a decision (Epstein et al. 1996). The taxonomy UI provided all the information and alternatives of design techniques at once, which corresponds to the internal information processing process of high rational decision-makers. This further leads to a fit between the tool for the selection task and the internal process of information processing. The corresponding fit results in a high perception of ease of use and lower perception of complexity of high rational decision-makers. On the contrary, low rational decision-makers had totally different perceptions of complexity and ease of use. For low rational decision-makers, too-much-information leads to more confusion instead of providing support (Epstein et al. 1996). Thus, the natural language UI comes with a higher perceived ease of use and lower perceived complexity than the taxonomy UI by the low rational decision-makers. Hence, under consideration of the rational decision-making style, a natural language UI is more helpful than a taxonomy UI for low rational decision-maker. Complementary, a taxonomy UI is more helpful than a natural language UI for high rational decision-maker.

7 Discussion[1]

In this chapter, the overall results of the dissertation are discussed. Specifically, the theoretical contributions, the practical contributions, limitations as well as future research are reflected. Table 7-1 presents a summary along the three cycles.

Table 7-1. A summary of key deliverables, contributions, limitations and future research

Cycle	Cycle 1	Cycle 2	Cycle 3
Key deliverables	Part 1: Expert-based taxonomy Part 2: Novice-based tags	Part 1: Initial design principles for a web-based platform Part 2: Influence of decision aids (taxonomy vs tags) on selection performance	Part 1: Refined design principles for an advisory platform Part 2: Influence of UIs (taxonomy vs natural language) on selection performance
Theoretical contributions	Part 1: Type I theory for analysis Part 2: Type I theory for analysis	Part 1: Type V theory for design and action Part 2: Type IV theory for explaining and predicting	Part 1: Type V theory for design and action Part 2: Type IV theory for explaining and predicting
Practical contributions	Part 1: A classification of design techniques from the experts' perspective Part 2: Capturing novices' categorization behavior using tags	Part 1: A web-based platform for supporting the selection of design techniques Part 2: Empirical Evidence that expert-based taxonomy is more helpful than tags for novice designers	Part 1: A web-based platform using different UIs for supporting the selection of design techniques Part 2: Empirical evidence that natural language UI fits rather short-term use; taxonomy-based UI suits rather for mid-/long-term use
Limitation & future research	Part 1: no relation between categories Part 2: some novices-created categories were merged	Part 1: evaluation focus on usability test without a performance test Part 2: borderline significant trend between the accuracy of tags and control	Part 1: no evaluation of long-term use Part 2: using an intelligent conversation agent instead of a natural language dialog
Outlook	i) A taxonomy of design techniques with overlapping characteristics under each dimension; ii) suggesting design techniques with packages for different design situations		

[1] This Chapter is based on the following studies which are published or in work: Liu et al. (2016); Liu, Leung, et al. (2018); Liu (2018); Liu, Werder, et al. (2018); Liu, He, et al. (2019); Liu, Werder, et al. (2019a); Liu, Werder, et al. (2019b); Liu, Rietz, et al. (2019)

7.1 Theoretical Contributions

Design Cycle 1

In part 1 of design cycle 1 a taxonomy of digital services design techniques has been suggested. Based on more than 300 data points, 70 design techniques were classified and provided the primary data for the development of the taxonomy. The suggested taxonomy results from the application of a systematic method (cf. Nickerson et al., 2013). It represents a theory for analysis (type I theory) (Gregor, 2006), which supports building a body of knowledge on design techniques in general. The rigorous developed dimensions and characteristics extend to existing research focusing on the classification of design techniques based on specific design situations (Maguire 2001). The study contributes a concise and extensible taxonomy with five dimensions and fifteen characteristics which are mutually exclusive and collectively exhaustive, following a systematic analysis of secondary data in conjunction with expert interviews. The taxonomy presents an overview of the similarities and differences of digital service design techniques, which provides a conceptual foundation that can be used as a theoretical basis for building selection support and the application of techniques. The identified dimensions help scholars to investigate antecedents of successful application of design techniques. A similar example can be found in the field of requirements engineering. Here, a methodology was developed based on the classified attributes of engineering techniques (Jiang et al. 2008). Furthermore, existing research used terminology inconsistently. While methods, techniques, and tools are clearly defined (cf. Brinkkemper, 1996), prior studies mix up these concepts regularly. Whereas studies that use different terms are referring to the same phenomenon, as well as using the same term referring to different phenomena. Such confusion is often referenced as a jingle-jangle fallacy (Larsen and Bong 2016). Given the different meaning of each term, they were clearly distinguished in this study. Researchers benefit from a clear distinction between methods, techniques, and tools when investigating the design process of digital services. In addition, a comprehensive list of unique design techniques was provided. It can be used in future studies to compare and contrast multiple design techniques.

In part 2 of design cycle 1, a novice-based classification of design techniques by conducting an open card sorting exercise was presented. The differences and similarities between novice- and expert-based classifications of design techniques were compared. In addition, differences of categorization behaviors between novices and experts in defining dimensions, creating categories, and assigning design techniques in specific categories were analyzed. The result shows that the actual behavior of categorizing design techniques is influenced by design knowledge and practical experience. Similar to part 1 of design cycle 1, the novice-based classification contributes in the form of a type I theory for analysis (Gregor 2006). This research advances the understanding on the differences and similarities between novices and experts in categorization behaviors in the context of design techniques (Chi et al. 1981; Schenk et al. 1998). In addition, the developed novice-based classification emphasizes the importance of considering novices' understanding of design techniques in providing selection support for deciding appropriate design techniques.

Design Cycle 2

In part 1 of design cycle 2, meta-requirements and design principles for a web-based platform to support the selection of design techniques were proposed. This represents a type V theory contribution for design and action (Gregor 2006). It gives prescriptive guidance for researchers to instantiate artifacts for classifying design techniques. The two design principles include the suggestion of using both expert-based and novice-based classification as decision aids in a web-based platform to help with the selection. Furthermore, users' suggestion of new design techniques with an attempt to improve the collection and classification is considered.

In part 2 of design cycle 2, a lab experiment was conducted to investigate the relationships between using decision aids (expert- and novice-based classifications) and the selection results in the context of novice designers selecting design techniques for specific task scenarios. The lab experiment result contributes to a type IV theory (i.e., EP theory) that explains and predicts the effect of using different decision aids on selection accuracy (Gregor 2006). At first, this research fills the gap that classifications are so ingrained (Bailey 1994) and no systematic evaluation of whether they can help with the selection of suitable design techniques for specific design situations is existing. The result turns out that both top-down taxonomies and bottom-up tags can help with the selection of design techniques. However, the use of expert-based taxonomy led to higher accuracy than novice-based tags. Second, this research also extends cognitive fit theory (Vessey 1991; Vessey and Galletta 1991) by explaining and predicting the task performance (i.e., selection accuracy) will improve not only by the match between the problem representation and the task but also by the decrease of system cognitive effort caused by the problem representation (i.e., decision aids). The result shows that the relationship between the use of an expert-based taxonomy and the selection accuracy is partially mediated by the system cognitive effort which connects the problem representation and the task performance introduced in cognitive fit theory (Vessey 1991; Vessey and Galletta 1991). Additionally, the results demonstrate that the decrease of system cognitive effort leads to the increase of selection accuracy, which contributes to the existing research that explains the relationship between problem representation and system cognitive effort (Hong et al. 2004). Third, the role of decision-making style is considered when using decision aids to select design techniques. The results show that a rational decision-making style has a certain trend towards significantly moderating the use of a bottom-up taxonomy to select appropriate design techniques.

Design Cycle 3

In part 1 of design cycle 3, an advisory platform for suggesting design techniques is presented. The proposed design principles contribute to design knowledge theoretically by guiding the development of a class of advisory systems with the specific purpose of suggesting design techniques, which can be seen as type V theory for design and action (Gregor 2006). The advisory platform uses a taxonomy and tags as a basis for suggesting design techniques, which contributes to the theoretical understanding of instantiating the taxonomy and tags for supporting the selection of design techniques. Compared with the design principles in design cycle 2, newly added design principles such as using natural language UI to instantiate the taxonomy extend existing solutions to solve the problem of suggesting appropriate design techniques for specific design situations. In addition, the design principle that suggests saving the advisory history based on design projects is an attempt to organize design techniques by different design projects, which can benefit the suggestion of package of design techniques for specific projects in the future.

In part 2 of design cycle 3, the influence of two different UIs (taxonomy vs. natural language) on the design novices' selection performance under consideration of rational decision-making style of design techniques was analyzed. The results contribute a type IV theory for explaining and predicting (Gregor 2006) and answer the "how" of the effect of different UIs on task performance (Whetten 1989). In addition, this experiment research also extends the understanding of the influence of using different UIs on task performance by considering rational decision-making style and repeated use of UIs in the context of decision-making task (Mennecke et al. 2006). The results reflect that the taxonomy UI outperformed the natural language UI in selection accuracy and perceived task performance. In the decision-making process, when considering rational decision-making style, UIs can influence the perceived complexity and ease of use. When considering the time used across the tasks in the experiment, the taxonomy UI was regarded as more appropriate than natural language UI for high rational decision-makers for Mid-/long-term use and vice versa.

7.2 Practical Contributions

Design Cycle 1

From a practical perspective, in part 1 of design cycle 1, the taxonomy shows differences and similarities of design techniques and provides a systematic overview of design techniques for practitioners. Additionally, the quality criteria for developing the taxonomy in combination with the experts' statements can help practitioners identify suitable design techniques, given varying situations (Kettinger et al. 1997). In part 2 of design cycle 1, the derived standardized categories can be used as a novice-based classification. One of the forms to present the flat categories can be tag clouds (Conroy et al. 2009; Sinclair and Cardew-Hall 2008).

Design Cycle 2

Practically, in part 1 of design cycle 2, the delivered web-based platform ServiceDesignKIT 1.0 helps design novices select suitable design techniques by using major, complementary filters; one is based on a classification that is created by experts (i.e., a top-down structured classification), and another is created by novices (i.e., bottom-up suggested tags). In addition, ServiceDesignKIT 1.0 provides basic descriptions of design techniques. Moreover, users can add new techniques, and suggest bottom-up tags describing design techniques. Users are also enabled to add comments about techniques to exchange knowledge and communicate with other users. The application of the top-down classification and bottom-up tags as a filter and the collection of bottom-up suggested tags and users' comments, which in turn can benefit the research community as a starting point to explore the effects of classifications on the selection support of design techniques. In part 2 of design cycle 2, the experiment research reflects that, in the selection process, the hierarchically structured taxonomy is more helpful than a set of tags built by a small number of novices. Although tags are built by novices, which might be easier for novices to understand, tags cannot support novices in the selection as good as the structured expert-based taxonomy. Thus, an expert-based taxonomy should be considered as the superior approach helping in the selection process. Tags can be used as a supplement such as the use of navigation and tag cloud on a website (Sinclair and Cardew-Hall 2008). The investigation of the role of cognitive effort shows that the development of decision aids should also consider the reduction of novices' system cognitive effort, which will in turn positively influence the selection process. Furthermore, a rational decision-making style might not positively influence the design-making process by using an expert-based taxonomy when the decision-makers have limited design technique knowledge. But if the decision aid can be so effective on selection task, the negative effect of decision-making style can be neglected. Thus, the expert-base decision aid needs to be developed as accurate as possible in order to neutralize the negative influence trend of design novices' rational decision-making style.

Design Cycle 3

ServiceDesignKIT 2.0 is a useful and publicly available tool for design novices to select design techniques in different design situations. The design technique library enables users to have an overview of design techniques and select design techniques. A natural language UI is provided to help design novices with limited background knowledge select design techniques without understanding the definitions of all the terms in the taxonomy. Furthermore, the platform provides users with an individual view that allows them to manage their preferred design technique sets. In addition, the design techniques saved in the personal page can be organized by design projects, which can further collect bundles of design techniques for a specific design project. In part 2 of cycle 3, the effect of taxonomy and natural language UI on the selection performance were evaluated. The result depicts that the taxonomy UI outperformed the natural language UI, even though the natural language UI explained the abstract terms in an easier way. With the visualization of the structured categories and the complete information of design techniques, although the use of a taxonomy needs time to understand the meaning of all the terms included in the taxonomy UI, the more times people use the taxonomy, the more skilled people become. Whereas, a natural language UI provides suggestions of design techniques at the end of the dialog, which helps beginners to quickly accomplish the task without requiring background knowledge. However, the natural language UI cannot benefit users in selecting design techniques with the increase of repeated use. Thus, a possible practical suggestion for designing the UIs for the selection of design techniques from the experiment result is the following: If a user needs to get some initial suggestion of design techniques for a design situation for one-time use, more benefits will be provided by using a natural language dialog than a taxonomy. However, when a user needs to select design techniques several times and wants to get a deeper understanding of the categories of design techniques, the taxonomy UI can provide more benefits.

7.3 Limitations and Future Research

Design Cycle 1

In part 1 of design cycle 1, several limitations are identified. First, the taxonomy may be perceived as a mere categorization exercise. The dimensions were suggested serving as sufficient conditions as a basis for the successful selection of design techniques. However, these conditions require further evaluation. Second, the dimensions and characteristics in the taxonomy cannot explain the relationships between each other and the priority order of the consideration of dimensions (Bailey 1994). However, a taxonomy may lead to a multitude of new relations between the identified concepts. Hence, more empirical research is needed in order to provide valid and reliable measurements and investigate the causal relations between the different concepts. In part 2 of design cycle 1, an open card sorting exercise was used to collect data and generate categories from the perspective of novices. In the open card sorting exercise, students who attended some foundational courses related to design with very limited design experience were regarded as novices. Open card sorting was conducted with 40 students. Comparing with existing open card sorting studies (e.g., Burnay 2016; Wentzel et al. 2016), a sample size of 40 can be considered as a reasonable sample size, but only including students to classify design techniques was still narrow. In the analysis process, many categories were derived from the raw data. Because it was impossible to compare the 110 categories that were directly derived from the sorting result to compare with the expert-based categories, the categories with the same names and similar descriptions were merged. Some data might get lost in the data-processing stage. As standardized categories were created, not all the exact labels that were derived by card sorting participants could be used. But as the purpose of this study does not include the analysis of the labels of categories, it is reasonable not to make the novices' created categories too complex to compare.

Design Cycle 2

In part 1 of design cycle 2, there were some drawbacks of the evaluation of ServiceDesignKIT 1.0. The results of both lab and pilot field evaluation rely on rather a small sample size. Extensive evaluation with larger sample size and diverse participants will help to create more accurate and solid results and more in-depth feedback. In the field evaluation, the filter quality received a positive evaluation, but there was no investigation on how people combined top-down classification and bottom-up tags when filtering design techniques. Moreover, people's intention to use needs to be further analyzed and increased. More features are needed to motivate novices to use the web platform. In addition, the differences between novices and experts when using classifications need further research, which will also benefit the collaboration between them. In additions, long-term use may reveal unforeseen issues that require further work to identify and resolve. Future research may extend the currently used selection support by recommending bundles of design techniques based on specific design projects and situations. In part 2 of design cycle 2, the experiment analysis showed there was a borderline significant trend between the selection accuracy of using tags and no decision aid. Moreover, the following tests on the moderating effect of intuitive decision-making style were not significant. One of the reasons could be that the tags which were created by a small sample size of students might not be accurate enough for the selection of design techniques. When more people participate in the tagging process, the tags may perform better than the results in this study (Sinclair and Cardew-Hall 2008). Furthermore, except for investigating the use of taxonomy and tags for selecting design techniques, the research on the effect of decision aids on the selection can be extended to other types of decision aids and other decision-making tasks.

Design Cycle 3

ServiceDesignKIT 2.0 leveraged natural language dialog and received positive feedback in the usability evaluation. However, the feedback was based on perceived usability, which did not reflect the actual selection performance by using a natural language dialog. Whether the platform can help with the selection of design techniques for specific design projects still needs long-term study. In addition, further use of the platform for collecting data such as specific design projects with corresponding design techniques can be an extension. In part 2 of design cycle 3, the different effect between the taxonomy and natural language UI on novices' selection performance was evaluated. In order to keep the suggested design techniques for specific selection conditions consistent between the two different UIs, the natural language UI was not designed as an intelligent adaptive conversational agent, but as a static dialog with pre-defined questions without any social cues. In the future, it may be investigated whether the use of an intelligent adaptive conversational agent can outperform the taxonomy should be analyzed.

7.4 Outlook

A Taxonomy of Design Techniques with Overlapping Characteristics

The followed taxonomy development method (Nickerson et al. 2013) emphasizes that characteristics need to be mutually exclusive and collectively exhaustive under each dimension. Thus, the developed taxonomy of design techniques was mutually exclusive and collectively exhaustive without any overlaps. While the methods, techniques, and tools have different definitions, ambiguities can result since some techniques could be attributed to multiple characteristics under the same dimension. In addition, in some cases, a design technique can have more than one characteristic. Take the heuristic evaluation as an example. It is assigned to the characteristic questionnaire because there are studies that use it as questionnaires (Kan Peng et al. 2004; Yusoh and Matayong 2017). But this technique is also used by some designers as desk analysis, and hence, some individuals may tend to attribute heuristic evaluation to desk analysis.

The suggestion of new characteristic can lead to the further evolution of the taxonomy over time. However, the definition of taxonomy does not allow overlaps between characteristics, which make the possibility of adding overlapped characteristics or assigning a design technique under more than one characteristic impossible. An alternative to the strict taxonomy could be that a design technique was assigned to more than one characteristics based on the different percent probability. Figure 7-1 presents an example.

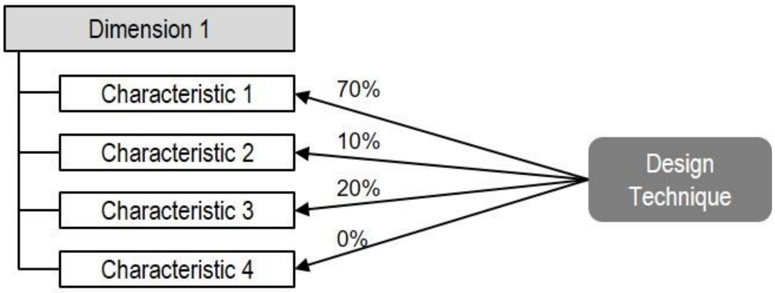


Figure 7-1. An example of taxonomy with overlapped characteristics

The overlapping taxonomy could benefit some situations more than the mutually exclusive and collectively exhaustive one (Nickerson et al. 2013). The suggestion of design techniques based on the overlapping taxonomy will not be focusing on one characteristic under a dimension but presenting the possibility of using a design technique under a design situation. The users of the overlapping taxonomy will have more design techniques to select after selecting basic dimensions and characteristics, which is followed by an in-depth selection based on the percent probability. However, the creation of such a taxonomy needs long-term data collection from a large sample size.

Suggesting Design Techniques with Packages for Different Design Situations

In practice, the application of design methods, techniques, and tools is rather ambiguous, which can be explained by the tight association of tools with techniques, and vice versa. For example, when creating a new mock-up, some designers may intuitively use their standard application, such as Balsamiq, and hence, use these terms almost synonymously. Therefore, practitioners can benefit from a configuration perspective toward design techniques (with its dimensions and characteristics) and their combinations with methods and tools. The interviews reflect these approaches in practice. In the interviews (Chapter 4.1), there were some statements that related to corresponding approaches described by experts: *"And after that is to think about different approaches to fulfill the needs, to create the software interface (Carl)."* Experts combined design methods and techniques when describing the activities for specific conditions. For example, Neil shared, *"We are also doing a lot of interviews with our own colleagues. […] And then for the different stories we are doing user stories. For the user stories, we also have personas."* In the first half of the quote, the expert mentioned interviews; while in the second half, two techniques (user stories and personas) were mentioned. The information could be summarized as follows: user stories are combined with personas, and the basic information for creating user stories and personas comes from interviews. In the design process, the activities do not merely rely on a specific design technique, but as a design approach that combines methods and techniques, which also correspond to the argument in Woolrych et al. (2011). The developed taxonomy in this study is based on the description of design techniques, which is able to provide an overview of a large number of design techniques and as a basis to help narrow down the search of design techniques. However, it does not yet consider the suggestion of different bundles of design techniques for specific project situations in combination with methods and tools. In the future, it would be interesting to suggest approaches that combine design methods, techniques, and tools as packages for specific projects.

8 Conclusion

As digital services are gradually complementing or even replacing traditional face-to-face services (Schneider 2017), more and more people from a multi-disciplinary background engage in the design process (Revang 2017). Before conducting design activities in the design process, suitable design techniques need to be selected. However, with the increasing number of design techniques, it is complex to choose the appropriate ones (Bollen et al. 2010; Greifeneder et al. 2010), especially for novice designers (Sanders et al. 2010).

With the awareness of the challenge of supporting design novices to select suitable design techniques for different design situations, the main research question of this dissertation project was derived, "*How to support design novices in digital service delivering organizations to select digital service design techniques?*" In order to answer this research question, three break-down research questions were formulated and answered step-by-step by structuring them into three design cycles following the DSR paradigm. In the first design cycle, two classifications of design techniques (i.e., a taxonomy and a set of tags) were created from the perspective of design experts and novices respectively (RQ1). In order to instantiate the two classifications in a tool to help with the selection, a web-based platform was developed, which included both classifications. To evaluate the different effect of taxonomy and tags on the novices' selection of design techniques, a lab experiment with a large sample size was conducted. The result showed that the taxonomy outperformed tags when design novices were selecting design techniques with a higher selection accuracy and a low system cognitive effort of using the classification (RQ2). Based on the results of design cycle 2, an advisor-based web platform was developed in design cycle 3, which used the taxonomy as the main support for the selection of design techniques with an attempt to additionally embed the taxonomy into a natural language dialog, which explained the dimensions and characteristics of the taxonomy in an easier way. The natural language UI and the original taxonomy UI were evaluated by a lab experiment with a large sample size. The results showed that the taxonomy UI, in general, outperformed the natural language UI in selection accuracy and perceived task performance. A deeper analysis of the time used in the tasks and the moderating effect of rational decision-making style on the relationship between different UIs and perceived complexity as well as perceived ease of use showed that the natural language UI was more appropriate for low rational decision-maker and the short-term use, whereas the taxonomy UI was more appropriate for high rational decision-maker and long-term use (RQ3).

To sum up, this dissertation contributes theoretically by providing classifications (i.e., a taxonomy and a set of tags) of design techniques and analysis of the effect of different decision aids on the selection performance. It also contributes practically by developing a web-bwsed platform as a tool to help design novices select design techniques based on different situations. I hope the findings of this research can serve as a reference for researchers and practitioners in the area of supporting design novices' selection of design techniques specifically and building classification systems in general.

9 References

Agarwal, R., and Venkatesh, V. 2002. “Assessing a Firm’s Web Presence: A Heuristic Evaluation Procedure for the Measurement of Usability,” *Information Systems Research* (13:2), pp. 168–186. (https://doi.org/10.1287/isre.13.2.168.84).

Agid, S. 2016. “‘...It’s Your Project, but It’s Not Necessarily Your Work...’: Infrastructuring, Situatedness, and Designing Relational Practice,” in *Proceedings of the 14th Participatory Design Conference on Full Papers - PDC ’16*, Aarhus, Denmark: ACM Press, pp. 81–90. (https://doi.org/10.1145/2940299.2940317).

Aguinis, H., and Gottfredson, R. K. 2010. “Best-Practice Recommendations for Estimating Interaction Effects Using Moderated Multiple Regression,” *Journal of Organizational Behavior* (31:6), pp. 776–786. (https://doi.org/10.1002/job.686).

Al-Wabil, A., Al-Husian, L., Al-Murshad, R., and Al-Nafjan, A. 2010. “Applying the Retrospective Think-Aloud Protocol in Usability Evaluations with Children: Seeing through Children’s Eyes,” in *2010 International Conference on User Science and Engineering (i-USEr)*, Shah Alam, Malaysia: IEEE, December, pp. 98–103. (https://doi.org/10.1109/IUSER.2010.5716731).

Alexander, P. A., Schallert, D. L., and Hare, V. C. 1991. “Coming to Terms: How Researchers in Learning and Literacy Talk About Knowledge,” *Review of Educational Research* (61:3), pp. 315–343. (https://doi.org/10.3102/00346543061003315).

Allinson, C. W., and Hayes, J. 1996. “The Cognitive Style Index: A Measure of Intuition-Analysis For Organizational Research,” *Journal of Management Studies* (33:1), pp. 119–135. (https://doi.org/10.1111/j.1467-6486.1996.tb00801.x).

Allison, D. 2012. “Chatbots in the Library: Is It Time?,” *Library Hi Tech* (30:1), pp. 95–107. (https://doi.org/10.1108/07378831211213238).

Alves, R., and Nunes, N. J. 2013. “Towards a Taxonomy of Service Design Methods and Tools,” in *Proceedings of the 4th International Conference on Exploring Services Science (IESS)*, J. F. e Cunha, M. Snene, and H. Nóvoa (eds.), Springer, Berlin, Heidelberg, pp. 215–229. (https://doi.org/10.1007/978-3-642-36356-6_16).

Armstrong, S. J. 2000. “The Influence of Individual Cognitive Style on Performance in Management Education,” *Educational Psychology* (20:3), pp. 323–339. (https://doi.org/10.1080/014434100750018020).

Arnold, V., Collier, P. A., Leech, S. A., and Sutton, S. G. 2004. “Impact of Intelligent Decision Aids on Expert and Novice Decision-Makers’ Judgments,” *Accounting and Finance* (44:1), pp. 1–26. (https://doi.org/10.1111/j.1467-629x.2004.00099.x).

Arora, S., Batra, K., and Singh, S. 2013. “Dialogue System: A Brief Review,” *ArXiv Preprint ArXiv:1306.4134*. (https://arxiv.org/ftp/arxiv/papers/1306/1306.4134.pdf).

Artz, M., Homburg, C., and Rajab, T. 2012. “Performance-Measurement System Design and Functional Strategic Decision Influence: The Role of Performance-Measure Properties,” *Accounting, Organizations and Society* (37:7), Elsevier Ltd, pp. 445–460. (https://doi.org/10.1016/j.aos.2012.07.001).

Asemi, A., Safari, A., and Asemi Zavareh, A. 2011. "The Role of Management Information System (MIS) and Decision Support System (DSS) for Manager's Decision Making Process," *International Journal of Business and Management* (6:7), pp. 164–173. (https://doi.org/10.5539/ijbm.v6n7p164).

Bailey, K. D. 1994. "Typologies and Taxonomies: An Introduction to Classification Techniques," *Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-102,* Thousand Oaks, CA: Sage.

Baker, J., Jones, D., and Burkman, J. 2009. "Using Visual Representations of Data to Enhance Sensemaking in Data Exploration Tasks," *Journal of the Association for Information Systems* (10:7), pp. 533–559. (https://doi.org/10.17705/1jais.00204).

Bangor, A., Kortum, P., and Miller, J. 2009. "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *Journal of Usability Studies* (4:3), pp. 114–123.

Bangor, A., Kortum, P. T., and Miller, J. T. 2008. "An Empirical Evaluation of the System Usability Scale," *International Journal of Human-Computer Interaction* (24:6), pp. 574–594. (https://doi.org/10.1080/10447310802205776).

Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology* (51:6), pp. 1173–1182. (https://doi.org/10.1037/0022-3514.51.6.1173).

Barrera-Barrera, R., Navarro-García, A., and Peris-Ortiz, M. 2015. "Evaluation of Quality in Different Electronic Services: A Multigroup Analysis," *International Journal of Management Science & Technology Information* (15), pp. 5–27.

Bettman, J. R., Johnson, E. J., and Payne, J. W. 1990. "A Componential Analysis of Cognitive Effort in Choice," *Organizational Behavior and Human Decision Processes* (45:1), pp. 111–139. (https://doi.org/10.1016/0749-5978(90)90007-V).

Bhattacherjee, A. 2012. *Social Science Research: Principles, Methods, and Practices*, (Second Edi.), Tampa, Florida: USF Tampa Bay Open Access Textbooks.

Blum, I. H., and Koskinen, P. S. 1991. "Repeated Reading: A Strategy for Enhancing Fluency and Fostering Expertise," *Theory Into Practice* (30:3), pp. 195–200. (https://doi.org/10.1080/00405849109543500).

Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H., and Winograd, T. 1977. "GUS, a Frame-Driven Dialog System," *Artificial Intelligence* (8:2), pp. 155–173. (https://doi.org/10.1016/0004-3702(77)90018-2).

Bollen, D., Knijnenburg, B. P., Willemsen, M. C., and Graus, M. 2010. "Understanding Choice Overload in Recommender Systems," in *Proceedings of the Fourth ACM Conference on Recommender Systems - RecSys '10*, Barcelona, Spain: ACM Press, pp. 63–70. (https://doi.org/10.1145/1864708.1864724).

Bonnardel, N., Lanzone, L., and Sumner, T. 2003. "Designing Web Sites: Opportunistic Actions and Cognitive Effort of Lay-Designers," *Cognitive Science Quarterly* (3:1), pp. 25–56.

Boulton, E., Hawley-Hague, H., Vereijken, B., Clifford, A., Guldemond, N., Pfeiffer, K., Hall, A., Chesani, F., Mellone, S., Bourke, A., and Todd, C. 2016. "Developing the FARSEEING

Taxonomy of Technologies: Classification and Description of Technology Use (Including ICT) in Falls Prevention Studies," *Journal of Biomedical Informatics* (61), Elsevier Inc., pp. 132–140. (https://doi.org/10.1016/j.jbi.2016.03.017).

Brhel, M., Meth, H., Maedche, A., and Werder, K. 2015. "Exploring Principles of User-Centered Agile Software Development: A Literature Review," *Information and Software Technology* (61), Elsevier B.V., pp. 163–181. (https://doi.org/10.1016/j.infsof.2015.01.004).

Brinkkemper, S. 1996. "Method Engineering: Engineering of Information Systems Development Methods and Tools," *Information and Software Technology* (38:4), pp. 275–280. (https://doi.org/10.1016/0950-5849(95)01059-9).

Broder, A. 2002. "A Taxonomy of Web Search," *ACM SIGIR Forum* (36:2), pp. 3–10. (https://doi.org/10.1145/792550.792552).

Brooke, J. 1996. "SUS - A Quick and Dirty Usability Scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland (eds.), London: Taylor & Francis, pp. 189–194.

Brucks, M. 1985. "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research* (12:1), pp. 1–16. (https://doi.org/10.1086/209031).

Brugnoli, G. 2011. "Connecting the Dots of User Experience," *Journal of Information Architecture* (1:1), pp. 6–15.

Bruun, A. 2010. "Training Software Developers in Usability Engineering," in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*, Reykjavik, Iceland: ACM Press, pp. 82–91. (https://doi.org/10.1145/1868914.1868928).

Buchanan, S., and McMenemy, D. 2012. "Digital Service Analysis and Design: The Role of Process Modelling," *International Journal of Information Management* (32), pp. 251–256. (https://doi.org/10.1016/j.ijinfomgt.2011.11.008).

Buchenau, M., and Suri, J. F. 2000. "Experience Prototyping," in *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, Brooklyn, New York, pp. 424–433. (https://doi.org/10.1145/347642.347802).

Burnay, C. 2016. "Are Stakeholders the Only Source of Information for Requirements Engineers ? Toward a Taxonomy of Elicitation Information Sources," *ACM Transactions on Management Information Systems* (7:3), 8:1-8:29. (https://doi.org/10.1145/2965085).

Burnette, C. H. 1982. "A Role Oriented Approach to Group Problem Solving," in *Group Planning and Problem Solving Methods in Engineering Management*, S. A. Olsen (ed.), New York: John Wiley and Sons, Inc., Chapter 10.

Burton-Jones, A., and Straub, D. W. 2006. "Reconceptualizing System Usage: An Approach and Empirical Test," *Information Systems Research* (17:3), pp. 228–246. (https://doi.org/10.1287/isre.1060.0096).

Bussolon, S. 2009. "Card Sorting, Category Validity, and Contextual Navigation," *Journal of Information Architecture* (1:2), pp. 5–29.

Cacioppo, J. T., Petty, R. E., and Feng Kao, C. 1984. "The Efficient Assessment of Need for

Cognition," *Journal of Personality Assessment* (48:3), pp. 306–307. (https://doi.org/10.1207/s15327752jpa4803_13).

Carayannopoulos, S. 2018. "Using Chatbots to Aid Transition," *International Journal of Information and Learning Technology* (35:2), pp. 118–129. (https://doi.org/10.1108/IJILT-10-2017-0097).

Cardoso, J., Fromm, H., Nickel, S., Satzger, G., Studer, R., and Weinhardt, C. 2015. *Fundamentals of Service Systems*, Service Science: Research and Innovations in the Service Economy, (J. Cardoso, H. Fromm, S. Nickel, G. Satzger, R. Studer, and C. Weinhardt, eds.), Cham: Springer International Publishing. (https://doi.org/10.1007/978-3-319-23195-2).

Chaturvedi, Dolk, and Drnevich. 2011. "Design Principles for Virtual Worlds," *MIS Quarterly* (35:3), p. 673. (https://doi.org/10.2307/23042803).

Chen, H., Liu, X., Yin, D., and Tang, J. 2017. "A Survey on Dialogue Systems: Recent Advances and New Frontiers," *ACM SIGKDD Explorations Newsletter* (19:2), pp. 25–35. (https://doi.org/10.1145/3166054.3166058).

Chi, M. T. H., Glaser, R., and Rees, E. 1981. "Expertise in Problem Solving," *Advances in the Psychology of Human Intelligence (1)*, (R. J. Sternberg, ed.), Hillsdale, NJ: Erlbaum.

Choi, Y. 2015. "A Complete Assessment of Tagging Quality: A Consolidated Methodology," *Journal of the Association for Information Science and Technology* (66:4), pp. 798–817. (https://doi.org/10.1002/asi.23198).

Churchill, G. A., and Surprenant, C. 1982. "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research* (19:4), pp. 491–504. (https://doi.org/10.1177/002224378201900410).

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Collier, J. E., and Bienstock, C. C. 2006. "Measuring Service Quality in E-Retailing," *Journal of Service Research* (8:3), pp. 260–275. (https://doi.org/10.1177/1094670505278867).

Collins, R. W., Hevner, A. R., Walton, G. H., and Linger, R. C. 2008. "The Impacts of Function Extraction Technology on Program Comprehension: A Controlled Experiment," *Information and Software Technology* (50:11), pp. 1165–1179. (https://doi.org/10.1016/j.infsof.2008.04.001).

Conroy, C., Brennan, R., O'Sullivan, D., and Lewis, D. 2009. "User Evaluation Study of a Tagging Approach to Semantic Mapping," in *The Semantic Web: Research and Applications: 6th European Semantic Web Conference, ESWC*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 623–637. (https://doi.org/10.1007/978-3-642-02121-3_46).

Corbin, J., and Strauss, A. 2014. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, (Fourth.), Thousand Oaks, CA: Sage.

Curedale, R. 2013. *Service Design: 250 Essential Methods*, Topanga CA: Design Community College Inc.

Curtis, G., and Vertelney, L. 1990. "Storyboards and Sketch Prototypes for Rapid Interface Visualization," in *CHI Tutorial*, ACM Press.

Davis, F. D. 1989. “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” *MIS Quarterly* (13:3), pp. 319–340. (https://doi.org/10.2307/249008).

Davis, F. D., and Yi, M. Y. 2004. “Improving Computer Skill Training: Behavior Modeling, Symbolic Mental Rehearsal, and the Role of Knowledge Structures,” *Journal of Applied Psychology* (89:3), pp. 509–523. (https://doi.org/10.1037/0021-9010.89.3.509).

Dell, N., and Kumar, N. 2016. “The Ins and Outs of HCI for Development,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16*, CHI ’16, San Jose, CA: ACM Press, pp. 2220–2232. (https://doi.org/10.1145/2858036.2858081).

Deloitte Access Economics. 2015. “Digital Government Transformation.” (http://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-digital-government-transformation-230715.pdf).

Deloitte Digital. 2019. “Embracing Acceleration in the Age of Digital Influence in Retail.” (https://www.deloittedigital.com/content/dam/deloittedigital/us/documents/blog/blog_20190226_google_mobile_influence_study.pdf).

DeSanctis, G. 1984. “Computer Graphics As Decision Aids: Directions for Research,” *Decision Sciences* (15:4), pp. 463–487. (https://doi.org/10.1111/j.1540-5915.1984.tb01236.x).

Druckman, J. N., and Kam, C. D. 2012. “Students as Experimental Participants,” in *Cambridge Handbook of Experimental Political Science*, J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia (eds.), Cambridge: Cambridge University Press, pp. 41–57. (https://doi.org/10.1017/CBO9780511921452.004).

Duchowski, A. T. 2002. “A Breadth-First Survey of Eye-Tracking Applications,” *Behavior Research Methods, Instruments, & Computers* (34:4), pp. 455–470. (https://doi.org/10.3758/BF03195475).

Dumais, S., Jeffries, R., Russell, D. M., Tang, D., and Teevan, J. 2014. “Understanding User Behavior Through Log Data and Analysis,” in *Ways of Knowing in HCI*, New York, NY: Springer New York, pp. 349–372. (https://doi.org/10.1007/978-1-4939-0378-8_14).

Dumas, J. S., and Salzman, M. C. 2006. “Usability Assessment Methods,” *Reviews of Human Factors and Ergonomics* (2:1), pp. 109–140. (https://doi.org/10.1177/1557234X0600200105).

Einhorn, H. J., and Hogarth, R. M. 1975. “Unit Weighting Schemes for Decision Making,” *Organizational Behavior and Human Performance* (13:2), pp. 171–192. (https://doi.org/10.1016/0030-5073(75)90044-6).

Elwyn, G., Frosch, D., Volandes, A. E., Edwards, A., and Montori, V. M. 2010. “Investing in Deliberation: A Definition and Classification of Decision Support Interventions for People Facing Difficult Health Decisions,” *Medical Decision Making* (30:6), pp. 701–711. (https://doi.org/10.1177/0272989X10386231).

Engin, A., and Vetschera, R. 2017. “Information Representation in Decision Making: The Impact of Cognitive Style and Depletion Effects,” *Decision Support Systems* (103), pp. 94–103. (https://doi.org/10.1016/j.dss.2017.09.007).

Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. 1996. “Individual Differences in Intuitive–

Experiential and Analytical–Rational Thinking Styles," *Journal of Personality and Social Psychology* (71:2), pp. 390–405. (https://doi.org/10.1037/0022-3514.71.2.390).

Erez, A., and Judge, T. A. 2001. "Relationship of Core Self-Evaluations to Goal Setting, Motivation, and Performance," *Journal of Applied Psychology* (86:6), pp. 1270–1279. (https://doi.org/10.1037//0021-9010.86.6.1270).

ESCO. 2017. "How to Implement ESCO." (https://ec.europa.eu/esco/portal/howtouse/6af39243-1e25-484b-af62-ad223a9c48b8).

Evenson, S. 2005. "Designing Design: Establishing a New Common Ground for Collaboration," in *HCI Session on Interaction Design Education and Research: Current and Future Trends. 11th International Conference on Human-Computer Interaction*, Las Vegas, NV.

Fan, Q., Yul Lee, J., and In Kim, J. 2013. "The Impact of Web Site Quality on Flow-related Online Shopping Behaviors in C2C E-marketplaces," *Managing Service Quality: An International Journal* (23:5), Emerald, pp. 364–387. (https://doi.org/10.1108/MSQ-11-2012-0150).

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences," *Behavior Research Methods* (39:2), pp. 175–191. (https://doi.org/10.3758/BF03193146).

Fawcett, T. 2006. "An Introduction to ROC Analysis," *Pattern Recognition Letters* (27:8), pp. 861–874. (https://doi.org/10.1016/j.patrec.2005.10.010).

Fitzgerald, T. 2013. "GUInform : Interactive Fiction for GUI Prototyping." (https://www.pdx.edu/computer-science/sites/www.pdx.edu.computer-science/files/Tesca Fitzgerald_0.pdf).

Flynn, L. R., and Goldsmith, R. E. 1999. "A Short, Reliable Measure of Subjective Knowledge," *Journal of Business Research* (46:1), pp. 57–66. (https://doi.org/10.1016/S0148-2963(98)00057-5).

Fonteyn, M. E., Kuipers, B., and Grobe, S. J. 1993. "A Description of Think Aloud Method and Protocol Analysis," *Qualitative Health Research* (3:4), pp. 430–441. (https://doi.org/10.1177/104973239300300403).

Fuge, M., Tee, K., Agogino, A., and Maton, N. 2014. "Analysis of Collaborative Design Networks: A Case Study of OpenIDEO," *Journal of Computing and Information Science in Engineering* (14:2), pp. 021009-1-021009-. (https://doi.org/10.1115/1.4026510).

Galletta, D. F., and Dunn, B. K. 2014. "Assessing Smartphone Ease of Use and Learning from the Perspective of Novice and Expert Users: Development and Illustration of Mobile Benchmark Tasks," *AIS Transactions on Human-Computer Interaction* (6:4), pp. 74–91. (https://aisel.aisnet.org/thci/vol6/iss4/1).

Garbarino, E. C., and Edell, J. A. 1997. "Cognitive Effort, Affect, and Choice," *Journal of Consumer Research* (24:2), pp. 147–158. (https://doi.org/10.1086/209500).

Gartner. 2019. "Customer Experience Design and Execution Overview." (https://emtemp.gcom.cloud/ngw/globalassets/en/marketing/documents/cx-design-and-execution.pdf).

Gerrard, V., and Sosa, R. 2014. "Examining Participation," in *Proceedings of the 13th*

Participatory Design Conference on Research Papers - PDC '14, Windhoek, Namibia, pp. 111–120. (https://doi.org/10.1145/2661435.2661451).

Glass, R. L., and Vessey, I. 1995. "Contemporary Application-Domain Taxonomies," *IEEE Software* (12:4), pp. 63–76. (https://doi.org/10.1109/52.391837).

Goh, D. H., Chua, A., Lee, C. S., and Razikin, K. 2009. "Resource Discovery through Social Tagging: A Classification and Content Analytic Approach," *Online Information Review* (33:3), pp. 568–583. (https://doi.org/10.1108/14684520910969961).

Golder, S. A., and Huberman, B. A. 2006. "Usage Patterns of Collaborative Tagging Systems," *Journal of Information Science* (32:2), pp. 198–208. (https://doi.org/10.1177/0165551506062337).

Gorlenko, L., and Englefied, P. 2006. "Usability Error Classification," in *CHI '06 Extended Abstracts on Human Factors in Computing Systems - CHI EA '06*, Montréal, Québec, Canada: ACM Press, pp. 803–808. (https://doi.org/10.1145/1125451.1125610).

Graf, B., Krüger, M., Müller, F., Ruhland, A., and Zech, A. 2015. "Nombot: Simplify Food Tracking," in *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia - MUM '15*, Linz, Austria: ACM Press, pp. 360–363. (https://doi.org/10.1145/2836041.2841208).

Gray, D., Brown, S., and Macanufo, J. 2010. "Capter 5: Games for Opening," in *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*, C. Wheeler (ed.), Sebastopol, CA: "O'Reilly Media, Inc., pp. 78–79.

Gregor, S. 2006. "The Nature of Theory in Information Systems," *MIS Quarterly* (30:3), pp. 611–642. (https://doi.org/10.2307/25148742).

Greifeneder, R., Scheibehenne, B., and Kleber, N. 2010. "Less May Be More When Choosing Is Difficult: Choice Complexity and Too Much Choice," *Acta Psychologica* (133:1), pp. 45–50. (https://doi.org/10.1016/j.actpsy.2009.08.005).

Gruber, M., de Leon, N., George, G., and Thompson, P. 2015. "Managing by Design," *Academy of Management Journal* (58:1), pp. 1–7. (https://doi.org/10.5465/amj.2015.4001).

van den Haak, M. J., de Jong, M. D. T., and Schellens, P. J. 2006. "Constructive Interaction: An Analysis of Verbal Interaction in a Usability Setting," *IEEE Transactions on Professional Communication* (49:4), pp. 311–324. (https://doi.org/10.1109/TPC.2006.885865).

Hackathorn, R. D., and Karimi, J. 1988. "A Framework for Comparing Information Engineering Methods," *MIS Quarterly* (12:2), p. 203. (https://doi.org/10.2307/248845).

Hamilton, K., Shih, S.-I., and Mohammed, S. 2016. "The Development and Validation of the Rational and Intuitive Decision Styles Scale," *Journal of Personality Assessment* (98:5), pp. 523–535. (https://doi.org/10.1080/00223891.2015.1132426).

Hassenzahl, M., and Ullrich, D. 2007. "To Do or Not to Do: Differences in User Experience and Retrospective Judgments Depending on the Presence or Absence of Instrumental Goals," *Interacting with Computers* (19:4), pp. 429–437. (https://doi.org/10.1016/j.intcom.2007.05.001).

Henrikson, R., De Araujo, B., Chevalier, F., Singh, K., and Balakrishnan, R. 2016. "Storeoboard:

Sketching Stereoscopic Storyboards," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI'16*, CHI '16, San Jose, California: ACM Press, pp. 4587–4598. (https://doi.org/10.1145/2858036.2858079).

Hepenstal, S., Kodagoda, N., Zhang, L., Paudyal, P., and Wong, B. L. W. 2019. "Algorithmic Transparency of Conversational Agents," in *Joint Proceedings Ofthe ACM IUI2019Workshops*, Los Angeles, USA, pp. 1–11. (http://ceur-ws.org/Vol-2327/IUI19WS-IUIATEC-1.pdf).

Hertzum, M., and Jacobsen, N. E. 1999. "The Evaluator Effect during First-Time Use of the Cognitive Walkthrough Technique," in *In Proceedings of the Eight International Conference on Human-Computer Interaction, (HCI International ´99),* Hillsdale, NJ, pp. 1063–1067.

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75–105. (https://doi.org/10.2307/25148625).

Hileman, R. 1998. "An Introductory Lecture for Digital Designers by Rhodes Hileman (c) 1998. [Online] Regarding: Jones, J. C., 1970. Design Methods: Seeds of Human Futures. New York & Chichester: John Wiley & Sons," New York & Chichester: John Wiley & Sons. (http://www.smsys.com/pub/dsgnmeth.pdf, accessed October 1, 2017).

Hong, W., Thong, J. Y. L., and Tam, K. Y. 2004. "The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective," *Journal of Management Information Systems* (21:3), pp. 149–184. (https://doi.org/10.1080/07421222.2004.11045812).

Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, J., and Baesens, B. 2011. "An Empirical Evaluation of the Comprehensibility of Decision Table, Tree and Rule Based Predictive Models," *Decision Support Systems* (51:1), pp. 141–154. (https://doi.org/10.1016/j.dss.2010.12.003).

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J., and Lehtinen, M. 2008. "Gaze Path Stimulation in Retrospective Think-Aloud," *Journal of Eye Movement Research* (2:4), pp. 1–18. (https://bop.unibe.ch/JEMR/article/view/2277/3473).

IDEO. 2015. *The Field Guide to Human-Centered Design*. (http://www.designkit.org//resources/1).

Iqbal, R., Grzywaczewski, A., Halloran, J., Doctor, F., and Iqbal, K. 2017. "Design Implications for Task-Specific Search Utilities for Retrieval and Re-Engineering of Code," *Enterprise Information Systems* (11:5), pp. 738–757. (https://doi.org/10.1080/17517575.2015.1086494).

Isomursu, M., Kuutti, K., and Väinämö, S. 2004. "Experience Clip: Method for User Participation and Evaluation of Mobile Concepts," in *Proceedings of the Eighth Conference on Participatory Design Artful Integration: Interweaving Media, Materials and Practices - PDC 04,* Toronto, Canada: ACM Press, pp. 83–92. (https://doi.org/10.1145/1011870.1011881).

Ives, B., Olson, M. H., and Baroudi, J. J. 1983. "The Measurement of User Information Satisfaction," *Communications of the ACM* (26:10), pp. 785–793. (https://doi.org/10.1145/358413.358430).

Iyengar, S. S., and Lepper, M. R. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?," *Journal of Personality and Social Psychology* (79:6), pp. 995–1006. (https://doi.org/10.1037/0022-3514.79.6.995).

Jeffiies, R., Turner, A. A., Poison, P., and Atwood, M. 1981. "The Processes Involved in Designing Software," in *Cognitive Skills and Their Acquisition*, J. R. Anderson (ed.), Hillsdale, NJ: Lawrence Erlbaum, pp. 255–283.

Jeffries, R., Miller, J. R., Wharton, C., and Uyeda, K. M. 1991. "User Interface Evaluation in the Real World: A Comparison of Four Techniques," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Reaching through Technology - CHI '91*, New Orleans, Louisiana: ACM Press, pp. 119–124. (https://doi.org/10.1145/108844.108862).

Jiang, L., Eberlein, A., Far, B. H., and Mousavi, M. 2008. "A Methodology for the Selection of Requirements Engineering Techniques," *Software & Systems Modeling* (7:3), pp. 303–328. (https://doi.org/10.1007/s10270-007-0055-y).

Johnson, E. J., and Payne, J. W. 1985. "Effort and Accuracy in Choice," *Management Science* (31:4), pp. 395–414. (https://doi.org/10.1287/mnsc.31.4.395).

Johnson, J. 2010. "Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules," *Morgan Kaufmann*. (http://cdsweb.cern.ch/record/1315413).

Judge, T. K., Pyla, P. S., McCrickard, D. S., Harrison, S., and Hartson, H. R. 2008. "Studying Group Decision Making in Affinity Diagramming," *Technical Report TR-08-16, Computer Science, Virginia Tech*. (http://eprints.cs.vt.edu/archive/00001043/).

Kan Peng, L., Ramaiah, C. K., and Foo, S. 2004. "Heuristic-based User Interface Evaluation at Nanyang Technological University in Singapore," *Program* (38:1), pp. 42–59. (https://doi.org/10.1108/00330330410519198).

Karimi, S., Papamichail, K. N., and Holland, C. P. 2015. "The Effect of Prior Knowledge and Decision-Making Style on the Online Purchase Decision-Making Process: A Typology of Consumer Shopping Behaviour," *Decision Support Systems* (77), Elsevier B.V., pp. 137–147. (https://doi.org/10.1016/j.dss.2015.06.004).

Keen, P. G. W. 1981. "Value Analysis: Justifying Decision Support Systems," *MIS Quarterly* (5:1), pp. 1–15. (https://doi.org/10.2307/249154).

Kettinger, W. J., Teng, J. T. C., and Guha, S. 1997. "Business Process Change: A Study of Methodologies, Techniques, and Tools," *MIS Quarterly* (21:1), pp. 55–80. (https://doi.org/10.2307/249742).

Kim, D. J., Ferrin, D. L., and Rao, H. R. 2009. "Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration," *Information Systems Research* (20:2), pp. 237–257. (https://doi.org/10.1287/isre.1080.0188).

Kim, J. H., Gunn, D. V, Schuh, E., Phillips, B., Pagulayan, R. J., and Wixon, D. 2008. "Tracking Real-Time User Experience (TRUE): A Comprehensive Instrumentation Solution for Complex Systems," in *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*, Florence, Italy: ACM Press, pp. 443–451. (https://doi.org/10.1145/1357054.1357126).

Kim, S. S., and Son, J. 2009. "Out of Dedication or Constraint? A Dual Model of Post-Adoption

Phenomena and Its Empirical Test in the Context of Online Services," *MIS Quarterly* (33:1), pp. 49–70. (https://doi.org/10.2307/20650278).

Kirk, C. P., Chiagouris, L., Lala, V., and Thomas, J. D. E. 2015. "How Do Digital Natives and Digital Immigrants Respond Differently to Interactivity Online? A Model for Predicting Consumer Attitudes and Intentions to Use Digital Information Products," *Journal of Advertising Research* (55:1), pp. 81–94. (https://doi.org/10.2501/JAR-55-1-081-094).

Kitchenham, B., Linkman, Stephen, and Linkman, Susan. 2005. "Experiences of Using an Evaluation Framework," *Information and Software Technology* (47:11), pp. 761–774. (https://doi.org/10.1016/j.infsof.2005.01.001).

Klašnja-Milićević, A., Vesin, B., Ivanović, M., Budimac, Z., and Jain, L. C. 2017. "Folksonomy and Tag-Based Recommender Systems in E-Learning Environments," in *E-Learning Systems* (Vol. 112), Cham: Springer, pp. 77–112. (https://doi.org/10.1007/978-3-319-41163-7_7).

Kohavi, R., Henne, R. M., and Sommerfield, D. 2007. "Practical Guide to Controlled Experiments on the Web," in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '07*, San Jose, California: ACM Press, pp. 959–967. (https://doi.org/10.1145/1281192.1281295).

Kolodner, J. L. 1983. "Towards an Understanding of the Role of Experience in the Evolution from Novice to Expert," *International Journal of Man-Machine Studies* (19:5), pp. 497–518. (https://doi.org/10.1016/S0020-7373(83)80068-6).

Koutsabasis, P., Spyrou, T., and Darzentas, J. S. 2007. "On the Performance of Novice Evaluators in Usability Evaluations," in *Proceedings of the 11th Panhellenic Conference on Informatics*.

Kozhevnikov, M. 2007. "Cognitive Styles in the Context of Modern Psychology: Toward an Integrated Framework of Cognitive Style," *Psychological Bulletin* (133:3), pp. 464–481. (https://doi.org/10.1037/0033-2909.133.3.464).

Krippendorff, K. 2004. *Content Analysis: An Introduction to Its Methodology*, (Second.), Thousand Oaks, CA: Sage.

Kuechler, W. L., and Vaishnavi, V. K. 2012. "A Framework for Theory Development in Design Science Research: Multiple Perspectives," *Journal of the Association for Information Systems* (13:6), pp. 395–423. (https://aisel.aisnet.org/jais/vol13/iss6/3).

Kumar, V., and LaConte, V. 2012. *101 Design Methods: A Structured Approach for Driving Innovation in Your Organization*, Hoboken, New Jersey: John Wiley & Sons, Inc.

Kutsikos, K., Konstantopoulos, N., Sakas, D., and Verginadis, Y. 2014. "Developing and Managing Digital Service Ecosystems: A Service Science Viewpoint," *Journal of Systems and Information Technology* (16:3), Emerald, pp. 233–248. (https://doi.org/10.1108/JSIT-02-2014-0015).

Larsen, K. R., and Bong, C. H. 2016. "A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses," *MIS Quarterly* (40:3), pp. 529–551. (https://doi.org/10.25300/MISQ/2016/40.3.01).

Lee, C.-C., Cheng, H. K., and Cheng, H.-H. 2007. "An Empirical Study of Mobile Commerce in

Insurance Industry: Task–Technology Fit and Individual Differences," *Decision Support Systems* (43:1), pp. 95–110. (https://doi.org/10.1016/j.dss.2005.05.008).

Lewis, J. R. 1994. "Sample Sizes for Usability Studies: Additional Considerations," *Human Factors: The Journal of the Human Factors and Ergonomics Society* (36:2), pp. 368–378. (https://doi.org/10.1177/001872089403600215).

Lewis, K. M., and Hepburn, P. 2010. "Open Card Sorting and Factor Analysis: A Usability Case Study," *The Electronic Library* (28:3), pp. 401–416. (https://doi.org/10.1108/02640471011051981).

Lilien, G. L., Rangaswamy, A., Van Bruggen, G. H., and Starke, K. 2004. "DSS Effectiveness in Marketing Resource Allocation Decisions: Reality vs. Perception," *Information Systems Research* (15:3), pp. 216–235. (https://doi.org/10.1287/isre.1040.0026).

Liu, X. 2018. "Selection Support of Digital Service Design Techniques for Design Novices," in *Proceedings of the Doctoral Consortium Papers at the 30th International Conference on Advanced Information Systems Engineering (CAiSE 2018)* (Vol. 2114), pp. 1–9. (http://ceur-ws.org/Vol-2114/paper1.pdf).

Liu, X., He, S., and Maedche, Al. 2019. "Designing an AI-Based Advisory Platform for Design Techniques," in *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm, Sweden, pp. 1–16. (https://aisel.aisnet.org/ecis2019_rp/152/).

Liu, X., Leung, E. T.-M., Toreini, P., and Maedche, A. 2018. "ServiceDesignKIT: A Web Platform of Digital Service Design Techniques," in *Proceedings of 13th International Conference on Design Science Research in Information Systems and Technology*, S. Chatterjee, K. Dutta, and R. Sundarraj (eds.), Cham: Springer, pp. 34–48. (https://doi.org/10.1007/978-3-319-91800-6_3).

Liu, X., Rietz, T., and Alexander, M. 2019. "Influence of UIs on the Novices' Selection Performance – A Comparison of Taxonomy and Natural Language Dialog."

Liu, X., Werder, K., and Mädche, A. 2016. "A Taxonomy of Digital Service Design Techniques," in *Proceedings of the 37th International Conference on Information Systems (ICIS)*, Dublin, pp. 1–12. (https://aisel.aisnet.org/icis2016/Human-ComputerInteraction/Presentations/5/).

Liu, X., Werder, K., and Maedche, A. 2019a. "A Taxonomy of Design Techniques for Digital Services."

Liu, X., Werder, K., and Maedche, A. 2019b. "Supporting Novice Designers' Decision-Making with Interactive Decision Aids – A Comparison of High and Low Structured Decision Aids."

Liu, X., Werder, K., and Zhao, Q. 2018. "The Role of Cultural Differences When Using Different Classifications: An Experiment Design for Design Technique Selection," in *Proceedings of the 39th International Conference on Information Systems (ICIS)*, San Francisco, pp. 1–9. (https://aisel.aisnet.org/icis2018/design/Presentations/1/).

Lurie, N. H., and Mason, C. H. 2007. "Visual Representation: Implications for Decision Making," *Journal of Marketing* (71:1), pp. 160–177. (https://doi.org/10.1509/jmkg.71.1.160).

Lutters, E., van Houten, F. J. A. M., Bernard, A., Mermoz, E., and Schutte, C. S. L. 2014. "Tools and Techniques for Product Design," *CIRP Annals* (63:2), pp. 607–630. (https://doi.org/10.1016/j.cirp.2014.05.010).

Macdonal, E. K., Wilson, H. N., and Konus, U. 2012. "Better Customer Insights - in Real Time," *Harvard Business Review* (90:9), pp. 102–108. (https://doi.org/10.1038/NNANO.2012.119).

Mackay, J. M., and Elam, J. J. 1992. "A Comparative Study of How Experts and Novices Use a Decision Aid to Solve Problems in Complex Knowledge Domains," *Information Systems Research* (3:2), pp. 150–172. (https://doi.org/10.1287/isre.3.2.150).

Maes, P. 1994. "Agents That Reduce Work and Information Overload," *Communications of the ACM* (37:7), pp. 31–40. (https://doi.org/10.1145/176789.176792).

Mager, B. 2009. "Service Design as an Emerging Field," in *Designing Services with Innovative Methods*, S. Miettinen and M. Koivisto (eds.), University of Art and Design Helsinki, pp. 28–43.

Maguire, M. 2001. "Methods to Support Human-Centred Design," *International Journal of Human-Computer Studies* (55), pp. 587–634. (https://doi.org/10.1006/ijhc.2001.0503).

Mangalaraj, G., Nerur, S., Mahapatra, R., and Price, K. H. 2014. "Distributed Cognition in Software Design: An Experimental Investigation of the Role of Design Patterns and Collaboration," *MIS Quarterly* (38:1), pp. 249–274. (https://doi.org/10.25300/MISQ/2014/38.1.12).

Marlow, C., Naaman, M., Boyd, D., and Davis, M. 2006. "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, to Read," in *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia - HYPERTEXT '06*, New York, New York, USA: ACM Press, p. 31. (https://doi.org/10.1145/1149941.1149949).

Martin, B., and Hanington, B. 2012. *Universal Methods of Design*, Beverly: Rockport Publisher.

Martin, L. P. 1998. "Cognitive-Style Inventory," in *The Pfeiffer Library Volume 8* (2nd ed.), Jossey-Bass/Pfeiffer. (http://home.snu.edu/~jsmith/library/body/v08.pdf).

Mathes, A. 2004. "Folksonomies - Cooperative Classification and Communication Through Shared Metadata." (http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html, accessed December 10, 2018).

McKay, J., Marshall, P., and Hirschheim, R. 2012. "The Design Construct in Information Systems Design Science," *Journal of Information Technology* (27:2), pp. 125–139. (https://doi.org/10.1057/jit.2012.5).

McKeithen, K. B., Reitman, J. S., Rueter, H. H., and Hirtle, S. C. 1981. "Knowledge Organization and Skill Differences in Computer Programmers," *Cognitive Psychology* (13:3), pp. 307–325. (https://doi.org/10.1016/0010-0285(81)90012-8).

McTear, M. F. 2002. "Spoken Dialogue Technology: Enabling the Conversational User Interface," *ACM Computing Surveys* (34:1), pp. 90–169. (https://doi.org/10.1145/505282.505285).

Mennecke, B. E., Crossland, M. D., and Killingsworth, B. L. 2006. "Is a Map More than a Picture? The Role of SDSS Technology, Subject Characteristics, and Problem Complexity on Map Reading and Problem Solving," *MIS Quarterly* (24:4), p. 601. (https://doi.org/10.2307/3250949).

Messier, W. E. J. 1995. “Research in and Development of Audit-Decisions Aids,” in *Judgment and Decision-Making Research in Accounting and Auditing*, Cambridge Series on Judgment and Decision Making, R. H. Ashton and A. H. Ashton (eds.), Cambridge: Cambridge University Press, pp. 207–228. (https://doi.org/10.1017/CBO9780511720420.010).

Meth, H., Mueller, B., and Maedche, A. 2015. “Designing a Requirement Mining System,” *Journal of the Association for Information Systems* (16:9), pp. 799–837. (https://aisel.aisnet.org/jais/vol16/iss9/2).

Miller, G. A. 1956. “The Magical Number Seven, plus or Minus Two: Some Limits on Our Capacity for Processing Information,” *Psychological Review* (63:2), pp. 81–97. (https://doi.org/10.1037/h0043158).

Mirani, R., and King, W. R. 1994. “Impacts of End-User and Information Center Characteristics on End-User Computing Support,” *Journal of Management Information Systems* (11:1), pp. 141–166. (https://doi.org/10.1080/07421222.1994.11518034).

Mitchell, J. B. 1989. “Current Theories on Expert and Novice Thinking: A Full Faculty Considers the Implications for Legal Education,” *Journal of Legal Education* (39), pp. 275–297.

Moritz, S. 2005. *Service Design: Practical Access to an Evolving Field*, Köln International School of Design.

Myers, M. D. 1997. “Qualitative Research in Information Systems,” *MIS Quarterly* (21:2), p. 241. (https://doi.org/10.2307/249422).

Myers, M. D. 2009. *Qualitative Research in Business & Management*, (1st ed.), London: Sage.

Nah, F. F., and Benbasat, I. 2004. “Knowledge-Based Support in a Group Decision Making Context: An Expert-Novice Comparison,” *Journal of the Association for Information Systems* (5:3), pp. 392–420. (https://aisel.aisnet.org/jais/vol5/iss3/5).

Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. “A Method for Taxonomy Development and Its Application in Information Systems,” *European Journal of Information Systems* (22:3), pp. 336–359. (https://doi.org/10.1057/ejis.2012.26).

Nielsen, J. 2005. “Putting A/B Testing in Its Place,” pp. 1–5. (http://www.useit.com/Alertbox/20050815.Html, accessed November 22, 2017).

Nielsen, J., and Pernice, K. 2009. *Eyetracking Web Usability*, Berkeley, CA: New Riders.

Ordóñez, L., and Benson, L. 1997. “Decisions under Time Pressure: How Time Constraint Affects Risky Decision Making,” *Organizational Behavior and Human Decision Processes* (71:2), pp. 121–140. (https://doi.org/10.1006/obhd.1997.2717).

Ozen Seneler, C., Basoglu, N., and Daim, T. U. 2010. “An Empirical Analysis of the Antecedents of Adoption of Online Services,” *Journal of Enterprise Information Management* (23:4), Emerald, pp. 417–438. (https://doi.org/10.1108/17410391011061753).

Pajala, T. 2019. “Explaining Choice Quality with Decision Style, Cognitive Reflection and Decision Environment,” *Journal of the Operational Research Society*, Taylor & Francis, pp. 1–15. (https://doi.org/10.1080/01605682.2018.1495994).

Palvia, P., and Nosek, J. T. 1993. “A Field Examination of System Life Cycle Techniques and Methodologies,” *Information & Management* (25:2), pp. 73–84.

(https://doi.org/10.1016/0378-7206(93)90049-Y).

Pan, J., and Yin, Z. 2015. "A Practice on Wayfinding System Design with Service Design Thinking," in *Marcus A. (Eds) Design, User Experience, and Usability: Interactive Experience Design. DUXU 2015*. (Vol. LNCS 9188), Cham: Springer, pp. 400–411. (https://doi.org/10.1007/978-3-319-20889-3_38).

Parsons, J., and Wand, Y. 2008. "Using Cognitive Principles to Guide Classification in Information Systems Modeling," *MIS Quarterly* (32:4), pp. 839–868. (https://doi.org/10.2307/25148874).

Patrício, L., Fisk, R. P., Falcão e Cunha, J., and Constantine, L. 2011. "Multilevel Service Design: From Customer Value Constellation to Service Experience Blueprinting," *Journal of Service Research* (14:2), pp. 180–200. (https://doi.org/10.1177/1094670511401901).

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. 2007. "A Design Science Research Methodology for Information Systems Research," *Journal of Management Information Systems* (24:3), pp. 45–77. (https://doi.org/10.2753/MIS0742-1222240302).

Pereira, R. E. 2000. "Optimizing Human-Computer Interaction for the Electronic Commerce Environment," *Journal of Electronic Commerce Research* (1:1), pp. 23–44. (https://pdfs.semanticscholar.org/dcac/50f9154991dcf3941fa911361f8bc04f13a3.pdf).

Peters, I., and Stock, W. G. 2008. "Folksonomy and Information Retrieval," *Proceedings of the American Society for Information Science and Technology* (44:1), pp. 1–28. (https://doi.org/10.1002/meet.1450440226).

Petrie, H., and Precious, J. 2010. "Measuring User Experience of Websites," in *Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '10*, Atlanta, Georgia: ACM Press, pp. 3673–3678. (https://doi.org/10.1145/1753846.1754037).

Phillips, S. D., Pazienza, N. J., and Ferrin, H. H. 1984. "Decision-Making Styles and Problem-Solving Appraisal," *Journal of Counseling Psychology* (31:4), pp. 497–502. (https://doi.org/10.1037/0022-0167.31.4.497).

Pine, B. J., and Gilmore, J. H. 1998. "Welcome to the Experience Economy," *Harvard Business Review* (July-Augus), pp. 97–105.

Prat, N., Comyn-Wattiau, I., and Akoka, J. 2015. "A Taxonomy of Evaluation Methods for Information Systems Artifacts," *Journal of Management Information Systems* (32:3), pp. 229–267. (https://doi.org/10.1080/07421222.2015.1099390).

Di Prospero, A., Norouzi, N., Fokaefs, M., and Litoiu, M. 2017. "Chatbots as Assistants: An Architectural Framework," in *Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering - CASCON '17*, Markham, Ontario, pp. 76–86. (https://dl.acm.org/citation.cfm?id=3172805).

Qahri-Saremi, H., Mueller-luckey, G., Robinson, R., and Sattovia, S. 2018. "Actualization of Electronic Health Records Affordances : An Empirical Investigation of Users ' Personal and Behavioral Antecedents," *Hicss*, pp. 3029–3038. (https://doi.org/10.24251/HICSS.2018.383).

Raymond, L. 1985. "Organizational Characteristics and MIS Success in the Context of Small

Business," *MIS Quarterly* (9:1), pp. 37–52. (https://doi.org/10.2307/249272).

Revang, M. 2017. "What Is the Most Important KPI in User Experience?" (https://blogs.gartner.com/magnus-revang/2017/10/03/what-is-the-most-important-kpi-in-user-experience/, accessed May 3, 2018).

Rieman, J., Franzke, M., and Redmiles, D. 1995. "Usability Evaluation with the Cognitive Walkthrough," in *Conference Companion on Human Factors in Computing Systems - CHI '95*, Denver, Colorado: ACM Press, pp. 387–388. (https://doi.org/10.1145/223355.223735).

Righi, C., James, J., Beasley, M., Day, D. L., Fox, J. E., Gieber, J., Howe, C., and Ruby, L. 2013. "Card Sort Analysis Best Practices," *Journal of Usability Studies* (8:3), pp. 69–89.

Robinson, M. D., Johnson, J. T., and Herndon, F. 1997. "Reaction Time and Assessments of Cognitive Effort as Predictors of Eyewitness Memory Accuracy and Confidence," *Journal of Applied Psychology* (82:3), pp. 416–425. (https://doi.org/10.1037/0021-9010.82.3.416).

Roschuni, C., Agogino, A. M., and Beckman, S. L. 2011. "The DesignExchange: Supporting the Design Community of Practice," in *Proceedings of the 18th International Conference on Engineering Design (ICED11)*, Denmark, pp. 255–264.

Roschuni, C., Kramer, J., and Agogino, A. 2015. "Design Talking: How Design Practitioners Talk About Design Research Methods," in *Proceedings of the 12th International Conference on Design Education*, Boston, Massachusetts: ASME, August 2, pp. 1–8. (https://doi.org/10.1115/DETC2015-47843).

Roschuni, C., Kramer, J., Qian, Z., Zakskorn, L., and Agogino, A. 2015. "Design Talking: An Ontology of Design Methods To Support a Common Language of Design," in *Proceedings of the 20th International Conference on Engineering Design (ICED 15)*, Milan, Italy, pp. 285–294.

Rugg, G., and McGeorge, P. 1997. "The Sorting Techniques: A Tutorial Paper on Card Sorts, Picture Sorts and Item Sorts," *Expert Systems* (14:2), pp. 80–93. (https://doi.org/10.1111/1468-0394.00045).

Sakata, J., Zhang, M., Pu, S., Xing, J., and Versha, K. 2014. "Beam: A Mobile Application to Improve Happiness and Mental Health," in *Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI EA '14*, CHI EA'14, New York, USA: ACM Press, pp. 221–226. (https://doi.org/10.1145/2559206.2580931).

Samuels, S. J. 1979. "The Method of Repeated Readings," *The Reading Teacher* (32:4), pp. 403–408. (https://www.jstor.org/stable/20194790).

Sanders, E. B.-N., Brandt, E., and Binder, T. 2010. "A Framework for Organizing the Tools and Techniques of Participatory Design," in *Proceedings of the 11th Biennial Participatory Design Conference on - PDC '10*, Sydney: ACM Press, pp. 195–198. (https://doi.org/10.1145/1900441.1900476).

Sangiorgi, U. B., Beuvens, F., and Vanderdonckt, J. 2012. "User Interface Design by Collaborative Sketching," in *Proceedings of the Designing Interactive Systems Conference on - DIS '12*, Newcastle: ACM Press, pp. 378–387. (https://doi.org/10.1145/2317956.2318013).

Schall, E., Ospina, S., Godsoe, B., and Dodge, J. 2004. "Appreciative Narratives as Leadership Research: Matching Method to Lens," in *Constructive Discourse and Human Organization*, D. L. Cooperrider and M. Avital (eds.), Burlington: Elsevier, pp. 147–170. (https://doi.org/10.1016/S1475-9152(04)01007-5).

Schenk, K. D., Vitalari, N. P., and Davis, K. S. 1998. "Differences between Novice and Expert Systems Analysts: What Do We Know and What Do We Do?," *Journal of Management Information Systems* (15:1), pp. 9–50. (https://doi.org/10.1080/07421222.1998.11518195).

Schleicher, D., Jones, P., and Kachur, O. 2010. "Bodystorming as Embodied Designing," *Interactions* (17:6), pp. 47–51. (https://doi.org/10.1145/1865245.1865256).

Schmettow, M., and Sommer, J. 2016. "Linking Card Sorting to Browsing Performance – Are Congruent Municipal Websites More Efficient to Use?," *Behaviour and Information Technology* (35:6), Taylor & Francis, pp. 452–470. (https://doi.org/10.1080/0144929X.2016.1157207).

Schneider, C. 2017. "10 Reasons Why AI-Powered, Automated Customer Service Is the Future." (https://www.ibm.com/blogs/watson/2017/10/10-reasons-ai-powered-automated-customer-service-future/, accessed January 11, 2018).

Schoenfeld, A. H., and Herrmann, D. J. 1982. "Problem Perception and Knowledge Structure in Expert and Novice Mathematical Problem Solvers," *Journal of Experimental Psychology: Learning, Memory, and Cognition* (8:5), pp. 484–494. (https://doi.org/10.1037/0278-7393.8.5.484).

Schultze, U., and Avital, M. 2011. "Designing Interviews to Generate Rich Data for Information Systems Research," *Information and Organization* (21:1), pp. 1–16. (https://doi.org/10.1016/j.infoandorg.2010.11.001).

Scott, S. G., and Bruce, R. A. 1995. "Decision-Making Style: The Development and Assessment of a New Measure," *Educational and Psychological Measurement* (55:5), pp. 818–831. (https://doi.org/10.1177/0013164495055005017).

Shaft, T. M., and Vessey, I. 2006. "The Role of Cognitive Fit in the Relationship between Software Comprehension and Modification," *MIS Quarterly* (30:1), pp. 29–55. (https://doi.org/10.2307/25148716).

Shafto, P., and Coley, J. D. 2003. "Development of Categorization and Reasoning in the Natural World: Novices to Experts, Naive Similarity to Ecological Knowledge," *Journal of Experimental Psychology: Learning, Memory, and Cognition* (29:4), pp. 641–649. (https://doi.org/10.1037/0278-7393.29.4.641).

Sharma, V. S., Mehra, R., and Kaulgud, V. 2017. "What Do Developers Want? An Advisor Approach for Developer Priorities," in *2017 IEEE/ACM 10th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, IEEE, May, pp. 78–81. (https://doi.org/10.1109/CHASE.2017.14).

Simon, F., and Usunier, J.-C. 2007. "Cognitive, Demographic, and Situational Determinants of Service Customer Preference for Personnel-in-Contact over Self-Service Technology," *International Journal of Research in Marketing* (24:2), pp. 163–173. (https://doi.org/10.1016/j.ijresmar.2006.11.004).

Simpson, G. G. 1961. *Principles of Animal Taxonomy*, New York: Columbia University Press.

Sinclair, J., and Cardew-Hall, M. 2008. "The Folksonomy Tag Cloud: When Is It Useful?," *Journal of Information Science* (34:1), pp. 15–29. (https://doi.org/10.1177/0165551506078083).

Spagnoletti, P., Resca, A., and Lee, G. 2015. "A Design Theory for Digital Platforms Supporting Online Communities: A Multiple Case Study," *Journal of Information Technology* (30:4), Nature Publishing Group, pp. 364–380. (https://doi.org/10.1057/jit.2014.37).

Spaulding, A., and Weber, J. S. 2009. "Usability Engineering Methods for Interactive Intelligent Systems," *AI Magazine* (30:4), pp. 41–47. (https://doi.org/10.1609/aimag.v30i4.2272).

Speier, C., and Morris, M. G. 2003. "The Influence of Query Interface Design on Decision-Making Performance," *MIS Quarterly* (27:3), pp. 397–423. (https://doi.org/10.2307/30036539).

Speier, C., Vessey, I., and Valacich, J. S. 2003. "The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance," *Decision Sciences* (34:4), pp. 771–797. (https://doi.org/10.1111/j.1540-5414.2003.02292.x).

Spencer, D. 2009. *Card Sorting: Designing Usable Categories*, (M. Justak, ed.), Brooklyn, New York: Louis Rosenfeld.

Spiteri, L. F. 2007. "The Structure and Form of Folksonomy Tags: The Road to the Public Library Catalog," *Information Technology and Libraries* (26:3), pp. 13–25. (https://doi.org/10.6017/ital.v26i3.3272).

Stickdorn, M., and Schneider, J. 2011. *This Is Service Design Thinking*, Amsterdam: BIS Publishers.

Szopinski, D., Schoormann, T., and Kundisch, D. 2019. "Because Your Taxonomy Is Worth It: Towards a Framework for Taxonomy Evaluation," in *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm & Uppsala, Sweden, pp. 1–20. (https://aisel.aisnet.org/ecis2019_rp/104).

Tähti, M., and Niemelä, M. 2006. "3E – Expressing Emotions and Experiences," in *Proceeding of WP9 Workshop on Innovative Approaches for Evaluating Affective Systems*, Stockholm, pp. 1–5.

Thawesaengskulthai, N., and Tannock, J. D. T. 2008. "A Decision Aid for Selecting Improvement Methodologies," *International Journal of Production Research* (46:23), pp. 6721–6737. (https://doi.org/10.1080/00207540802230553).

Thompson, R. L., Higgins, C. A., and Howell, J. M. 1994. "Influence of Experience on Personal Computer Utilization: Testing a Conceptual Model," *Journal of Management Information Systems* (11:1), pp. 167–187. (https://doi.org/10.1080/07421222.1994.11518035).

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. 2014. "Mediation: R Package for Causal Mediation Analysis.," *Journal of Statistical Software* (59:5), pp. 1–38. (https://doi.org/10.18637/jss.v059.i05).

Todd, P., and Benbasat, I. 1999. "Evaluating Effort , and of DSS , Impact Cognitive Incentives on Strategy Selection," *Information Systems Research* (10:4), pp. 356–374. (https://doi.org/10.1287/isre.10.4.356).

Tohidi, M., Buxton, W., Baecker, R., and Sellen, A. 2006. "User Sketches: A Quick, Inexpensive, and Effective Way to Elicit More Reflective User Feedback," in *Proceedings of the 4th Nordic Conference on Human-Computer Interaction Changing Roles - NordiCHI '06*, Oslo: ACM Press, pp. 105–114. (https://doi.org/10.1145/1182475.1182487).

Tremblay, M. C., Hevner, A. R., and Berndt, D. J. 2010. "The Use of Focus Groups in Design Science Research," in *Communications of the Association for Information Systems* (Vol. 26), pp. 121–143. (https://doi.org/10.1007/978-1-4419-5653-8_10).

Tuch, A. N., Bargas-Avila, J. A., Opwis, K., and Wilhelm, F. H. 2009. "Visual Complexity of Websites: Effects on Users' Experience, Physiology, Performance, and Memory," *International Journal of Human Computer Studies* (67:9), Elsevier, pp. 703–715. (https://doi.org/10.1016/j.ijhcs.2009.04.002).

Turban, E., and Aronson, J. E. 2001. *Decision Support Systems and Intelligent Systems*, (6th ed.), Upper Saddle River, NJ: Prentice Hall.

Vaishnavi, V., and Kuechler, B. 2004. *Design Science Research in Information Systems*. (http://www.desrist.org/design-research-in-information-systems/).

Vanattenhoven, J., and De Roeck, D. 2010. "Investigating the Use of the Experience Clip Method," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering* (Vol. 40 LNICST), pp. 79–86. (https://doi.org/10.1007/978-3-642-12630-7_9).

Vargo, S. L., and Lusch, R. F. 2004. "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing* (68:1), pp. 1–17. (https://doi.org/10.1509/jmkg.68.1.1.24036).

Vermeeren, A. P. O. S., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J., and Väänänen-Vainio-Mattila, K. 2010. "User Experience Evaluation Methods: Current State and Development Needs," in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*, Reykjavik, Iceland: ACM Press, pp. 521–530. (https://doi.org/10.1145/1868914.1868973).

Vessey, I. 1985. "Expertise in Debugging Computer Programs: A Process Analysis," *International Journal of Man-Machine Studies* (23:5), pp. 459–494. (https://doi.org/10.1016/S0020-7373(85)80054-7).

Vessey, I. 1991. "Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature," *Decision Sciences* (22:2), pp. 219–240. (https://doi.org/10.1111/j.1540-5915.1991.tb00344.x).

Vessey, I., and Galletta, D. 1991. "Cognitive Fit: An Empirical Study of Information Acquisition," *Information Systems Research* (2:1), pp. 63–84. (https://doi.org/10.1287/isre.2.1.63).

Vigo, M., Bail, S., Jay, C., and Stevens, R. 2014. "Overcoming the Pitfalls of Ontology Authoring: Strategies and Implications for Tool Design," *International Journal of Human-Computer Studies* (72:12), Elsevier, pp. 835–845. (https://doi.org/10.1016/j.ijhcs.2014.07.005).

Virzi, R. A. 1992. "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?," *Human Factors: The Journal of the Human Factors and Ergonomics Society* (34:4), pp. 457–468. (https://doi.org/10.1177/001872089203400407).

Voss, J. 2007. “Tagging, Folksonomy & Co - Renaissance of Manual Indexing?,” in *10th International Symposium for Information Science*, Cologne, January 10, pp. 1–12. (http://arxiv.org/abs/cs/0701072).

Vries, G. de, Hartevelt, M., and Oosterholt, R. 1995. “Private Camera Conversation: A New Method for Eliciting User Responses?,” *Behaviour & Information Technology* (14:6), pp. 358–360. (https://doi.org/10.1080/01449299508914654).

Wand, Y., Monarchi, D. E., Parsons, J., and Woo, C. C. 1995. “Theoretical Foundations for Conceptual Modelling in Information Systems Development,” *Decision Support Systems* (15:4), pp. 285–304. (https://doi.org/10.1016/0167-9236(94)00043-6).

Wang, T., and Ji, P. 2010. “Understanding Customer Needs through Quantitative Analysis of Kano’s Model,” *International Journal of Quality & Reliability Management* (27:2), (B. Clegg, ed.), pp. 173–184. (https://doi.org/10.1108/02656711011014294).

Wang, W., and Benbasat, I. 2009. “Interactive Decision Aids for Consumer Decision Making in E-Commerce: The Influence of Perceived Strategy Restrictiveness,” *MIS Quarterly* (33:2), pp. 293–320. (https://doi.org/10.2307/20650293).

Wentzel, J., Müller, F., Beerlage-de Jong, N., and van Gemert-Pijnen, J. 2016. “Card Sorting to Evaluate the Robustness of the Information Architecture of a Protocol Website,” *International Journal of Medical Informatics* (86), pp. 71–81. (https://doi.org/10.1016/j.ijmedinf.2015.12.003).

Werder, K., Zobel, B., and Maedche, A. 2016. “PDISC – Towards a Method for Software Product DISCovery,” in *Software Business: 7th International Conference, ICSOB 2016, Ljubljana, Slovenia, June 13-14, 2016, Proceedings*, A. Maglyas and A.-L. Lamprecht (eds.), Cham: Springer International Publishing, pp. 47–62. (https://doi.org/10.1007/978-3-319-40515-5_4).

Whetten, D. A. 1989. “What Constitutes a Theoretical Contribution?,” *Academy of Management Review* (14:4), pp. 490–495. (https://doi.org/10.5465/amr.1989.4308371).

Williams, K., Chatterjee, S., and Rossi, M. 2008. “Design of Emerging Digital Services: A Taxonomy,” *European Journal of Information Systems* (17:5), pp. 505–517. (https://doi.org/10.1057/ejis.2008.38).

Wixom, B. H., and Todd, P. A. 2005. “A Theoretical Integration of User Satisfaction and Technology Acceptance,” *Information Systems Research* (16:1), pp. 85–102. (https://doi.org/10.1287/isre.1050.0042).

Wodehouse, A., and Ion, W. 2012. “Augmenting the 6-3-5 Method with Design Information,” *Research in Engineering Design* (23:1), pp. 5–15. (https://doi.org/10.1007/s00163-011-0110-0).

Woolrych, A., Hornbæk, K., Frøkjær, E., and Cockton, G. 2011. “Ingredients and Meals Rather Than Recipes: A Proposal for Research That Does Not Treat Usability Evaluation Methods as Indivisible Wholes,” *International Journal of Human-Computer Interaction* (27:10), pp. 940–970. (https://doi.org/10.1080/10447318.2011.555314).

Wulf, D., and Bertsch, V. 2017. “A Natural Language Generation Approach to Support Understanding and Traceability of Multi-Dimensional Preferential Sensitivity Analysis in Multi-Criteria Decision Making,” *Expert Systems with Applications* (83), Elsevier Ltd, pp.

131–144. (https://doi.org/10.1016/j.eswa.2017.04.041).

Xiao, B., and Benbasat, I. 2007. “E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact,” *MIS Quarterly* (31:1), pp. 137–209. (https://doi.org/10.2307/25148784).

Xu, J., Benbasat, I., and Cenfetelli, R. 2011. “The Effects of Service and Consumer Product Knowledge on Online Customer Loyalty,” *Journal of the Association for Information Systems* (12:11), pp. 741–766. (https://doi.org/10.17705/1jais.00279).

Yi, C., Jiang, Z. (Jack), and Benbasat, I. 2015. “Enticing and Engaging Consumers via Online Product Presentations: The Effects of Restricted Interaction Design,” *Journal of Management Information Systems* (31:4), pp. 213–242. (https://doi.org/10.1080/07421222.2014.1001270).

Yusoh, S., and Matayong, S. 2017. “Heuristic Evaluation of Online Satisfaction Survey System for Public Healthcare Service: Applying Analytical Hierarchical Process,” in *2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE, November, pp. 462–467. (https://doi.org/10.1109/ICITISEE.2017.8285551).

Zugal, S., and Pinggera, J. 2014. “Low–Cost Eye–Trackers: Useful for Information Systems Research?,” in *Advanced Information Systems Engineering Workshops*, L. Iliadis, M. Papazoglou, and K. Pohl (eds.), Cham: Springer International Publishing, pp. 159–170. (https://doi.org/10.1007/978-3-319-07869-4_14).

Appendices

Appendix A. Initial version of the taxonomy

Table Appendix-1. Initial version of the taxonomy

Dimensions	*Characteristics*	*Descriptions*
Design phases	Planning	Design techniques are usually applied for generating ideas or conducting user research.
	Draft prototyping	Design techniques support prototyping processes and comparison of a series of prototypes to select one for making detail prototypes.
	Detailed prototyping	Design techniques focus on evaluating the design work in detail to reach a stable version.
	Launching	Design techniques concentrate on gathering user feedback to improve digital services and to prepare for the next iteration.
Time dependency	Real-time	Design techniques are applied to reveal various users' immediate emotional feedback.
	Retrospective	Design techniques indicate the dynamics of emotion based on users' impressions of the experience.
Duration	Long-term study	Data need to be collected and analyzed over a long period.
	Short-term study	Design techniques can be used in a short period.
Participants	User involved	Real users need to be involved in the design process.
	Without user	Real users are not available in the design process.
Evaluation types	Questionnaire	Questionnaires indicate goal achievement, user satisfaction, etc.
	Interview	Interviews generate data to reveal users' experience and prospects.
	Experiment	Experiments reflect user behavior by means of concrete data.
	Observation	Designers observe users to understand user behavior.
	Group discussion	Innovations can be generated by collecting ideas from members of a discussion group in which a series of ideas can be evaluated.

Appendix B. Interview questions for evaluating the taxonomy

The interview included two parts. In the first part, questions on the design process and the used design techniques were asked without presenting our taxonomy. In the second part, the taxonomy was presented and experts' feedback was collected.

First Part:

The interview questions focused on evaluating the content of the taxonomy. As the ultimate goal of the taxonomy is to help with the design process, instead of directly asking questions such as "based on what dimensions and characteristics you have considered to select design techniques," questions in the context of design projects and design processes were asked. Main questions could be sorted into three groups.

1. Questions on a design project

The first questions were "Could you please describe a recent design project you have participated in?" and "How did you participate in this project?" All interviews started with these two questions because experts were expected to explain the used situations of design techniques based on a project that they had participated. Additionally, we could make sure that the described design situations related to design processes of digital services. The project represented the context that helped to describe the design techniques and was used in further questions as a reference point.

2. Questions on used design techniques

After experts described a recent design project, the next question was "What techniques have you used during the design process?" In the interview, experts described design techniques with combinations of methods (e.g., fast prototyping), techniques (e.g., wireframing), and tools (e.g., SketchFlow) during the design process. If an expert only explained a general method, such as prototyping, further questions were asked. For example: "What kind of prototypes did you create?" and "What techniques did you use when prototyping?" If an expert only mentioned the name of a technique and no detailed descriptions, a laddering approach to questioning was applied, for example, "Why and how did you apply the technique?"

3. Questions on design situations (i.e., dimensions and characteristics)

The term "situations" was used instead of "dimensions and characteristics" in the questions to avoid mentioning any information of the taxonomy in the first part of the interview. In addition, the term “situations” fitted well into the context of the described project and design techniques. For example, “What kind of situations have you considered when using the techniques?” and “Why did you consider these situations when using [mentioned technique]?” Some supplement questions included: “Who participated in the creation of prototypes?” and “Did you evaluate the prototype before moving to further development activities?”

Second Part:

At the beginning of part two, the interviewer presented and explained the taxonomy, and then started the interview. The following questions aimed at collecting experts’ feedback on the taxonomy.

1. Questions on dimensions and characteristics in the taxonomy

“Have you considered the dimensions and characteristics in the taxonomy when deciding what techniques to use in the design process?” This question was asked to see whether experts agreed/disagreed with the dimensions and characteristics in the taxonomy.

2. Questions on the usefulness of the taxonomy

“How do you think of the taxonomy?” “Do you think this taxonomy gives an overview of design techniques and can help you with choosing design techniques?” “Do you have any suggestions for the improvement of the taxonomy?”

Appendix C. Coding frequency and representative quotations

Table Appendix-2. Coding frequency and representative quotations

Dimensions & Characteristics	Inter-views	% Inter-views	Representative quotations
Design phase	*15*	*100%*	
Planning	15	100%	*"We first study the work practices they have currently." (Frank)* *"We are working on a project start with user research, and I am collecting first ideas." (Harry)* *"Normally, I start with the specification to know what we want for software design." (Irene)*
Draft prototyping	14	93%	*"I made paper prototyping and my colleagues need to play with the paper prototypes." (Grace)* *"And we did paper prototyping and photograph with that present the prototypes in PPT as a sort of pre-limited mock." (Max)* *"Then I start creating ideas… Sometimes, we just do paper prototype." (Neil)*
Detailed prototyping	14	93%	*"After this prototyping, users use the software prototype, and they will say 'OK. This may be better if in another way.'" (Bert)* *"For example, my colleagues from customer lab analyze the test and give us product management tips on how to concept a new feature or how to design the UI." (Emma)* *"When you already have a prototype and observe how users use your software." (Harry)*
Launching	5	40%	*"At the end, we release the App and try to follow some approaches, which directly deliver some feedback which is really close to us we can really get more information over a long period of time." (David)* *"And when the software is in the market, we get feedback from users. This will also be analyzed to get a deeper view of our users." (Emma)* *"That is now moving into commercial products that really brings both products that replace our pilot then." (Frank)*
Time-dependency	*12*	*80%*	
Real-time	11	73%	*"One other thing is visit on-site, we make the observation, pictures, recording everything, asking everything." (Bert)* *"We are working with the craftsman through the App so we have a better and detailed idea about the work they should do." (David)* *"We observe the users and look at how they solve their problems, then we look at how can we solve the problem in a software." (Emma)*
Retrospective	7	47%	*"The user told me "OK, I have to get the information on this and this" just without the App." (Alice)* *"We do the retrospective test." (Neil)*
Duration	*12*	*80.0%*	

Long-term	4	27%	*"We use 5-10 users to get feedback over the time." (David)* *"Or if there are more money, we can again make some iterations and improve the results." (Bert)* *"And in one or two years, we have the final products." (Oliver)*
Short-term	11	73%	*"I'd like to get really early feedback from the real craftsman in the early stage." (David)* *"We do more short-term study." (Harry)* *"But often, we are doing quick development." (Neil)*
Participant	*15*	*100%*	
User involved	14	93%	*"We tried to get insight from the users if the prototype needs that. ... We use a focus group to invite a lot of users, we asked about the processes and problems, the topic." (Bert)* *"So I actually invited policemen to come to my home and observe the working around my home because it is a little more difficult to go to other houses. We talked to them what they currently do, tried to get as much involved what they do to understand what their work is and what their practices and so on and so on." (Frank)* *"At the University, we talked to other students and asked if they are willing to participate." (Kevin)*
Without user	11	73%	*"So we at first writing down the facts, maybe creating some personas to give some ideas about the persona who's using it and describe the context." (David)* *"But we couldn't ask the user how we should go further." (Max)* *"The heuristic questionnaire is usually done by experts in this domain." (Emma)*
Evaluation types	*15*	*100%*	
Questionnaire	4	27%	*"I have used customized questionnaires to get some specific information about the context and the usage, but currently probably not usability questionnaires." (David)* *"After mock-up, I would say using a questionnaire." (Irene)* *"I take, for example, NASA TLX questionnaire." (Lisa)*
Interview	11	73%	*"We try to ask users' goals and situations to get more information in context with the user. This interview could be by phone or directly with users." (Bert)* *"I would say for the specification would be an interview because you talk very freely with the user." (Irene)* *"So we tried to do user interviews. We are also doing a lot of interviews with our own colleagues." (Neil)*
Experiment	5	33%	*"Sometimes, we also use a method like eye-tracking when we test Apps." (Emma)* *"First of all, experiments, we can control the behavior that we can make sure the weakness of our software doesn't go through the practice." (Frank)* *"The advantage of this Sketchflow is you can also test this prototype in a usability lab. You can also test in a usability lab at the end." (Jan)*
Observation	12	80%	*"So I start with the observation of five users vary from experienced users to who just start with their job." (Alice)*

			"We are looking for users' needs through observation. […] Then we observe the users and try to use think-aloud method to know what they think and about their problems when using our software." (Emma) *"I was sitting there, looking at them, and making my notes." (Grace)*
Group discussion	13	87%	*"So when you already have a product, then you can test it and show it to the product owner and team members and discuss on real approaches to see like the stories like these we have these steps, we have these buttons, and these text fields, and the stuff like that." (Carl)* *"So we can create some ideas and discuss the ideas in our small team." (David)* *"After that, we normally sit together with experts around and discuss how to build a prototype." (Emma)*

Appendix D. A taxonomy of digital service design techniques

Table Appendix-3. A taxonomy of design techniques for digital services

Dimensions and characteristics / Design techniques	*Design Phases*	Planning	Low-fidelity prototyping	High-fidelity prototyping	Release	*Time dependency*	Real-time	Retrospective	*Duration*	Long-term study	Short-term study	*User participation*	User attendance	User absence	*Evaluation types*	Questionnaire	Interview	Experiment	Observation	Group discussion
3-12-3 Brainstorming		x					x				x			x						x
3E (Expressing Experiences and Emotions)				x				x			x		x			x				
6-3-5 Brainwriting		x					x				x			x						x
A/B Testing				x			x				x		x					x		
Activity Map		x					x				x			x						x
Actors Mapping		x						x			x			x						x
Affinity Diagramming		x						x			x			x						x
Attribute Listing					x			x			x			x					x	
Behavioral Mapping					x		x			x			x						x	
Bodystorming			x				x				x			x						x
Business Origami			x				x				x			x						x
Co-Discovery				x			x				x		x						x	
Cognitive Mapping		x					x				x			x						x
Cognitive Walkthrough			x				x				x			x			x			
Collaborative Sketching			x				x				x			x						x
Concept Mapping		x					x				x		x							x
Concurrent Think-Aloud				x			x				x		x						x	
Content Inventory & Audit					x			x			x			x					x	
Contextual Laddering		x					x				x		x				x			
Critical Incident Technique					x		x				x		x				x			
Desirability Testing with Product Reaction Cards			x				x				x		x							x
Diary Studies					x			x		x			x			x				
Directed Storytelling		x					x				x			x						x
Experience Clips					x		x				x		x						x	
Experience Prototyping			x				x				x			x					x	
Eye-tracking				x			x				x		x					x		
Flexible Modeling			x				x				x		x							x
Fly-on-the-Wall Observation		x					x				x		x						x	
Graffiti Walls		x					x			x			x						x	
Greeting Cards		x					x				x			x						x
Heuristic Evaluation			x				x				x			x		x				
Kano Analysis		x					x				x		x			x				
LEGO Serious Play		x					x				x			x						x
Mental Model Diagramming		x						x		x			x							x
Mind Mapping		x						x			x			x						x
Mood Boards			x					x			x			x						x
Motivation Matrix		x						x			x			x		x				
Offering Mapping		x						x			x			x						x

Parallel prototyping		x			x		x			x					x
Personas	x					x		x		x					x
Phillips 66 Method	x				x			x		x					x
Photo Diary				x	x		x		x					x	
Photo Elicitation Interviewing	x					x	x		x			x			
Private Camera Conversation			x			x		x	x			x			
Product Experience Tracker				x		x	x		x		x				
Repertory Grids		x				x		x	x				x		
Retrospective Think-Aloud			x			x		x	x					x	
Role-Playing	x				x			x		x					x
Scenarios	x					x		x		x					x
Service Blueprints	x					x		x	x						x
Shadowing	x				x			x	x					x	
Speed Dating		x			x			x	x					x	
Stakeholder Walkthrough		x			x			x	x						x
Story Sharing	x				x			x		x					x
Storyboarding	x					x		x		x					x
Sustainability Map	x				x			x		x					x
Teachback		x			x			x	x			x			
Territory Maps	x				x			x		x					x
The Love Letter & the Breakdown Letter	x					x		x	x		x				
Thought Leader	x				x			x		x					x
Time Machine	x				x			x		x					x
Tomorrow Headlines	x				x			x		x					x
Touchpoint Matrix	x					x		x		x				x	
Trigger	x				x			x		x					x
User Journey Maps		x				x		x	x					x	
UX Curve				x		x	x		x		x				
Web Analytics				x	x		x		x					x	
Wireframing		x				x		x		x					x
Wishful Thinking	x				x			x		x					x
Wizard of Oz		x			x			x	x				x		

Appendix E. Analysis of card sorting result

Table Appendix-4. Overlaps between standardized categories and design techniques.

Card name	User Research	Idea generation	Information Organization	Feedback Collection	Prototype evaluation	Prototyping	Evaluation	Expert participation	Product Evaluation	Collaboration with stakeholders	User participation	Short term duration	Long term duration	UX Evaluation	Mid-term duration	Relationships and Dependencies
3-12-3 Brainstorming	5%	45%	10%				5%	15%		5%		10%				
3E- Expressing Emotion and Experience	26%			21%	16%		5%					16%		5%		5%
6-3-5 Brainwriting	5%	45%	10%				5%	15%		5%		10%				
A/B Testing	16%			16%	16%	5%	16%						5%		5%	
Actors Mapping	20%	10%	30%							5%	5%	5%	5%		5%	10%
Affinity Diagramming		40%	25%					15%					5%		5%	5%
Attribute Listing	10%	10%	20%	5%			5%	15%	5%			10%				
Behavioral Mapping	40%		20%	5%			5%						5%	5%	5%	5%
Bodystorming		16%		5%	16%	5%	5%	16%				11%		11%		5%
Business Origami	5%	40%						20%		10%		10%				5%
Closed Card Sorting	15%	25%	15%	5%			5%	5%				15%				5%
Co-discovery	30%			10%	10%		10%				5%	5%	5%	5%	5%	5%
Cognitive Mapping		25%	35%	5%	5%			10%				5%	5%			
Cognitive Walkthrough	5%	5%	11%		11%	16%	11%	21%				5%		5%		
Collaborative Sketching	11%	11%			5%	21%				5%	16%	16%				5%
Concept Mapping		30%	35%	5%	5%			10%			5%	10%				
Content Inventory & Auditing	5%		20%	10%		5%	10%	5%	5%			5%	5%			
Contextual Laddering	35%	5%		15%	5%						10%	15%				5%
Critical Incident Technique	32%	11%	5%				5%	5%		5%	5%		11%			5%
Concurrent Think-aloud	16%			21%	5%	5%	16%				11%	16%				5%
Desirability Testing with Product Research Cards	15%	5%	5%	20%		5%	10%	5%			5%	5%	10%	5%		5%
Diary Studies	42%	11%		5%				11%					16%			5%
Directed Storytelling	47%	11%					5%				11%	11%			5%	5%
Experience Clip	30%		5%	10%	5%		5%	5%				5%	5%	5%		5%
Experience Prototyping	5%	5%			16%	26%					5%	11%		11%	5%	
Eye-tracking	40%			5%			10%		10%		5%	10%				
Flexible Modeling	21%	5%	5%	5%	5%	21%					11%	16%				
Fly-on-the-Wall Observation	55%	10%						5%					5%		5%	5%
Graffiti Walls	25%	20%		10%			5%			5%	5%	10%	5%			5%
Heuristic Evaluation	11%		5%	16%	21%	5%	5%	21%				5%	5%			
Kano Analysis	35%	5%	5%		10%		10%					10%	5%			5%
LEGO Serious Play	5%	26%	5%			11%		11%		5%		16%		11%		
Mental Model Diagrams	10%	25%	20%				15%		5%		5%		5%			
Mind Mapping	5%	40%	25%					10%				10%				5%
Mood Boards	30%	20%	10%					5%			10%	15%				
Motivation Matrix	5%	25%	25%	5%				5%		5%		15%				5%
Offering Map	10%	20%	30%			5%		10%				10%		5%		5%
Open Card Sorting	25%	20%	10%	5%			5%				5%	10%			5%	5%

Parallel Prototyping	26%	5%	5%		11%	21%		5%					11%		5%	
Personas	20%	10%	15%		10%			15%					5%	10%	5%	5%
Photo Diary	40%	10%		5%				5%			10%	5%	10%			5%
Photo Elicitation Interview	45%	10%		5%							10%	10%			5%	5%
Private Camera Conversation	21%			16%	21%		11%					11%				5%
Product Experience Tracker	15%		5%	25%	5%		10%		10%		5%	5%	10%			
Repertory Grid Technique	16%		21%	16%	11%		11%					11%			5%	
Retrospective Think-aloud	11%			26%	11%		11%				16%	11%				5%
Roadmapping		30%	30%		5%			15%					5%		5%	
Role-playing	30%	10%			5%			15%				5%		10%		5%
Scenarios	30%	15%	10%		5%			10%		5%		10%		5%		5%
Sentence Completion	25%	5%		15%	5%		5%				10%	10%		5%	5%	5%
Service Blueprint	30%	5%	20%		5%	5%						10%	5%			
Shadowing	50%	10%		5%							5%		10%			5%
Speed Dating	15%			5%	15%	5%	10%		5%		10%	10%			5%	
Stakeholder Maps	5%	25%	25%	5%				5%		5%	5%	10%				10%
Stakeholder Walkthrough	5%	5%		5%	11%	5%	11%			16%	11%	16%		5%		5%
Storyboards	20%	35%	10%		5%			10%				10%				5%
Story Sharing	15%	25%				5%		20%		5%		10%				5%
Territory Maps		30%	25%					15%				10%				
The Love Letter & Breakup Letter	25%	5%		15%	5%		5%				5%	15%		5%		5%
Tomorrow Headlines	25%	30%	5%		10%							15%				
Touchpoints Matrix	25%	5%	20%	5%	5%			10%				10%		5%		10%
User Journey Maps	25%	5%	25%		5%						5%		5%			5%
UX Curve	10%	5%	15%	20%			10%		5%		5%		15%			
Value Mapping		20%	30%			5%		15%				10%				5%
Value Opportunity Analysis	5%	21%		5%			11%	16%	5%			11%				
Web Analytics	15%	5%	15%	10%			15%	5%	5%				10%			
Web of Abstraction	5%	30%	20%		5%	5%		10%				10%				
Weighted Matrix		15%	25%				10%	15%	5%	5%		10%				5%
Wireframe	5%	5%	16%		5%	26%		11%			5%	5%			5%	
Wizard of Oz	26%			11%	16%	11%	5%				5%	5%			11%	5%

Note: The percentage means the agreement of all participants on assigning a technique under a standardized category (Spencer 2009).

Appendix F. Tasks for the lab-based usability test for ServiceDesignKIT 1.0

Part 1: Read the scenario and use ServiceDesignKIT for tasks 0-3 (screen recording)

Part 2: Watch the recording from part 1 and in parallel speak out what you were thinking about when using the web tool (E.g., I was thinking…, when I clicked on this button/these buttons.) (voice recording)

Scenario:

- You plan to develop a new service in a specific area in a short period of time.
- You want to have an understanding of users before you start to create prototypes.

Task 0:

- Please use the following Testuser account to log in before starting the following 3 tasks.
  Email: testuser@sdk.com password: password

Task 1:

- You want to select up to 3 design techniques and add them to favorites.

Task 2:

- You want to have a look at the top discussed techniques.
- You want to add a short comment to one of the most discussed techniques with the sentence: "This is a helpful technique!"

Task 3:

- You want to submit a new technique with the following attributes (pay attention to markup, lists etc.) to the platform.

  Name: Technique1

  Description: Technique

  Instruction: 1. Technique

  Design Phase: Launching

  Time Dependency Retrospective

  Duration: Long Term

  User Participation: Without Real User

  Add tag: Useful

Appendix G. Three task scenarios and design techniques used in the experiment

Please imagine, you are working in a team on a software development project. Your project develops a mobile App enabling users to book cinema tickets online. For developing the App, you need to go through three stages. In the task, you need to select appropriate design techniques for each stage.

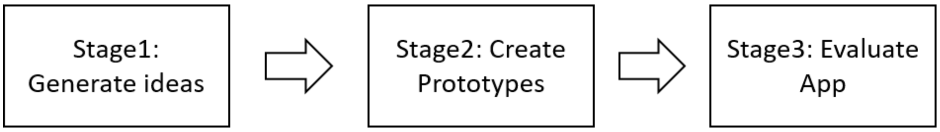


Figure Appendix-1. Experiment task scenarios

Table Appendix-5. Experiment task scenarios and correct selections

Task scenario	Design technique	Reference
Stage 1: Generate ideas • You are at an early stage of your project. • You want to involve a group of people (more than two) to discuss the project and generate some ideas during the discussion. • You do not have access to real users.	3-12-3 Brainstorming	(Gray et al. 2010)
	6-3-5 Brainwriting	(Wodehouse and Ion 2012)
	Role-playing	(Burnette 1982)
	Story sharing	(IDEO 2015)
	Tomorrow headlines	(Moritz 2005)
	Offering mapping	(Pan and Yin 2015)
	Storyboarding	(Curtis and Vertelney 1990)
Stage 2: Create prototypes • You now want to create some initial prototypes based on the ideas created in stage 1. • After creating several prototypes, you want to compare them and decide on one or two prototypes for further refinement. • You again do not have access to real users when creating and evaluating prototypes.	Bodystorming	(Buchenau and Suri 2000)
	Experience prototyping	(Buchenau and Suri 2000)
	Heuristic evaluation	(Agarwal and Venkatesh 2002)
	Collaborative sketching	(Sangiorgi et al. 2012)
	Cognitive walkthrough	(Rieman et al. 1995)
	Wireframing	(Fitzgerald 2013)
	Mood board	(Buchenau and Suri 2000)
Stage 3: Evaluate App • You now have developed a first running version of the app	Concurrent think-aloud	(Petrie and Precious 2010)
	Retrospective think-aloud	(Hyrskykari et al. 2008)

• You want to evaluate the app with real users before delivering it on a large scale to the market • The evaluation seeks to collect real users' data and feedback within a couple of days.	Private camera conversation	(Vries et al. 1995)
	A/B testing	(Nielsen 2005)
	Co-discovery	(Koutsabasis et al. 2007)
	Product experience tracker	(Macdonal et al. 2012)
	Experience clips	(Isomursu et al. 2004)

Table Appendix-6. Design techniques used in the experiment

3-12-3 Brainstorming	Experience Clips	Role-Playing
6-3-5 Brainwriting	Experience Prototyping	Scenarios
A/B Testing	Heuristic Evaluation	Service Blueprints
Actors Mapping	Kano Analysis	Shadowing
Behavioral Mapping	Mental Model Diagramming	Stakeholder Walkthrough
Bodystorming	Mood Boards	Story Sharing
Co-Discovery	Motivation Matrix	Storyboarding
Cognitive Walkthrough	Offering Mapping	Tomorrow Headlines
Collaborative Sketching	Personas	Touchpoint Matrix
Concurrent Think-Aloud	Photo Elicitation Interviewing	User Journey Maps
Desirability Testing with Product Reaction Cards	Private Camera Conversation	UX Curve
Diary Study	Product Experience Tracker	Wireframing
Directed Storytelling	Repertory Grids	Wizard of Oz
	Retrospective Think-Aloud	

Appendix H. Screenshot for the experiment training session (Chapter 5.2)

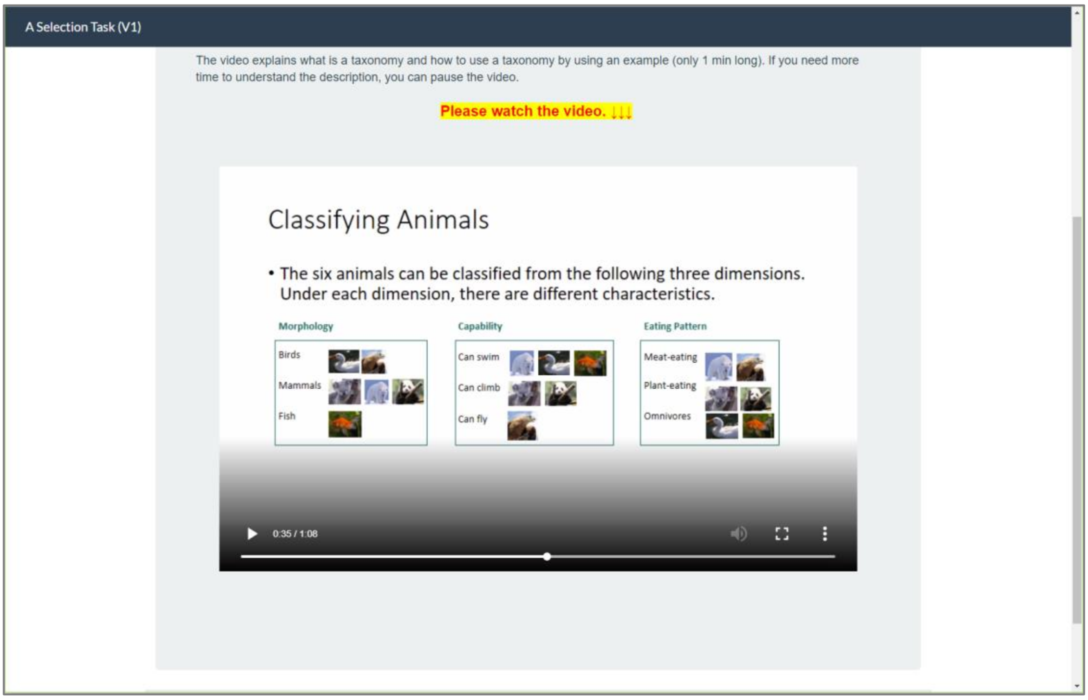


Figure Appendix-2. Screenshot of the description of creating a taxonomy

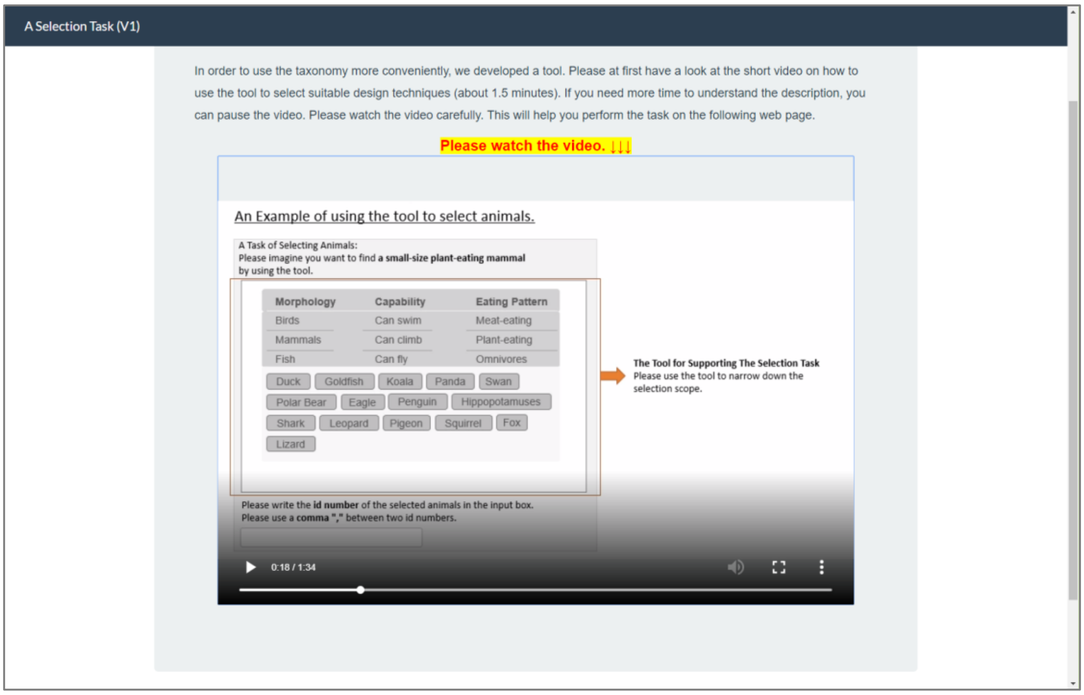


Figure Appendix-3. Screenshot of the description of using a taxonomy

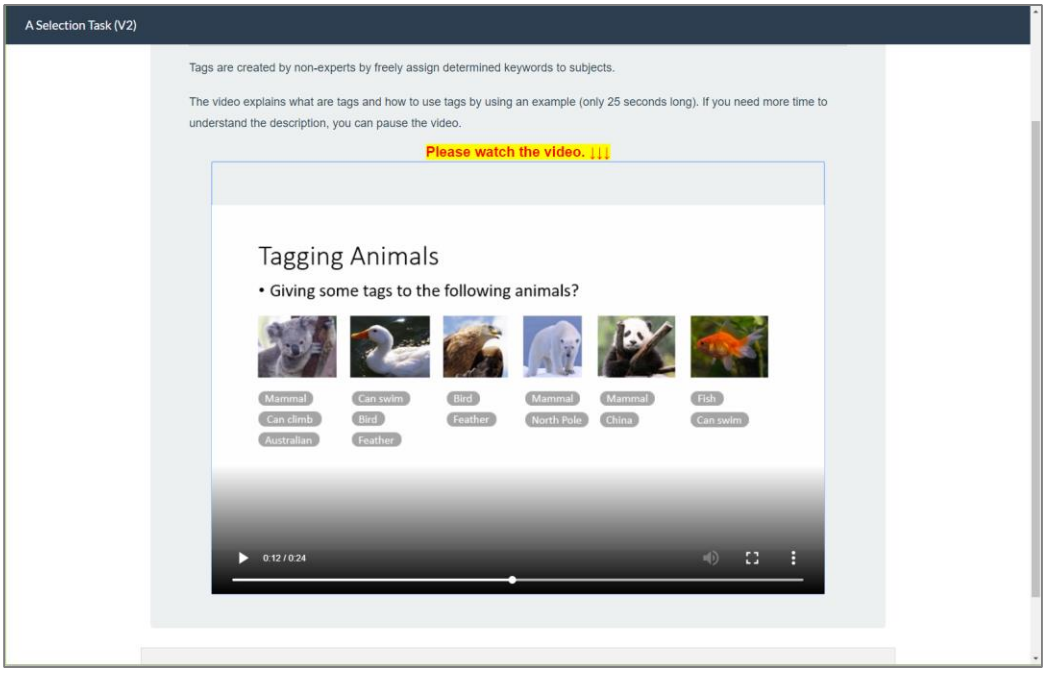


Figure Appendix-4. Screenshot of the description of creating tags

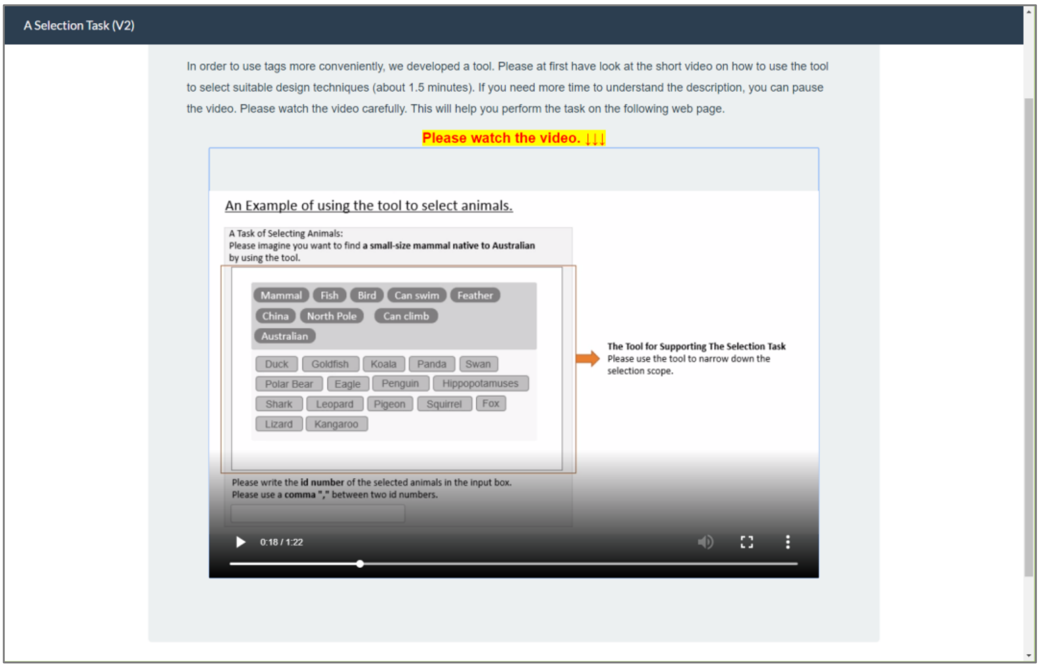


Figure Appendix-5. Screenshot of the description of using tags

Appendix I. Measurements and items used for evaluating the effect of taxonomy and tags on selection

Table Appendix-7. Measurements and items used in Chapter 5.2

Measurement	Item
Decision-making Styles (Hamilton et al. 2016)	Rational decision-making style 1. I prefer to gather all the necessary information before coming to a decision. 2. I thoroughly evaluate decision alternatives before making a final choice. 3. In decision making, I take time to contemplate the pros/cons or risks/benefits of a situation. 4. Investigating the facts is an important part of my decision-making process. 5. I weigh a number of different factors when making decisions. Intuitive decision-making style 1. When making decisions, I rely mainly on my gut feelings. 2. My initial hunch about decisions is generally what I follow. 3. I make decisions based on intuition. 4. I rely on my first impressions when making decisions. 5. I weigh feelings more than analysis in making decisions.
System Cognitive Effort (Pereira 2000)	1. To complete the task, using [the taxonomy/tags/list] was very frustrating. 2. To complete the task, using [the taxonomy/tags/list] took too much time. 3. To complete the task, using [the taxonomy/tags/list] required too much effort. 4. To complete the task, using [the taxonomy/tags/list] was too complex. 5. To complete the task, using [the taxonomy/tags/list] was easy.
Perceived Design Technique Knowledge (Davis and Yi 2004; Flynn and Goldsmith 1999)	1. I know pretty much about design techniques. 2. In a design group, I am one of the experts when using design techniques. 3. Compare to most other members in a design team, I know less about these products. 4. When using a design technique, I really do not know a lot. 5. I do not feel very knowledgeable about design techniques. 6. I have a lot of experiences with design techniques. 7. I feel familiar with design techniques.

Appendix J. Rotated factor loading (Chapter 5.2)

Table Appendix-8. Rotated factor loading (Chapter 5.2)

Item	RC 1	RC 2	RC 3	RC 4	h2	u2
System cognitive effort 01	**0.75**	-0.15	0	-0.01	0.58	0.42
System cognitive effort 02	**0.86**	0.03	0	-0.11	0.75	0.25
System cognitive effort 03	**0.89**	-0.08	0.09	-0.04	0.8	0.2
System cognitive effort 04	**0.83**	-0.07	0.14	0.08	0.72	0.28
System cognitive effort 05	**0.82**	-0.03	0.03	-0.01	0.68	0.32
Rational decision-making 01	-0.06	**0.78**	-0.21	0.11	0.68	0.32
Rational decision-making 02	-0.08	**0.83**	-0.1	0.12	0.71	0.29
Rational decision-making 03	-0.02	**0.82**	-0.07	0.05	0.68	0.32
Rational decision-making 04	-0.11	**0.76**	-0.19	0.01	0.63	0.37
Rational decision-making 05	-0.03	**0.73**	-0.12	-0.12	0.57	0.43
Intuitive decision-making 01	0	-0.11	**0.87**	0.03	0.77	0.23
Intuitive decision-making 02	0.08	-0.11	**0.86**	0.09	0.77	0.23
Intuitive decision-making 03	-0.03	-0.07	**0.85**	0.01	0.73	0.27
Intuitive decision-making 04	0.13	-0.17	**0.79**	0.09	0.67	0.33
Intuitive decision-making 05	0.11	-0.31	**0.65**	0.16	0.55	0.45
Perceived technique knowledge 01	0	0.07	0.06	**0.87**	0.77	0.23
Perceived technique knowledge 02	0.06	0.04	0.14	**0.88**	0.79	0.21
Perceived technique knowledge 03	0.03	-0.04	0.06	**0.76**	0.58	0.42
Perceived technique knowledge 04	-0.06	-0.11	0.03	**0.72**	0.53	0.47
Perceived technique knowledge 05	0	0.03	0.02	**0.87**	0.75	0.25
Perceived technique knowledge 06	-0.05	0.1	0.03	**0.77**	0.6	0.4
Perceived technique knowledge 07	-0.08	0.08	0.05	**0.81**	0.67	0.33

Notes: Extraction methods: principal component factoring; rotation method: varimax rotation; h2: share of variance explained by factors; u2: share of variance not explained by factors.

Appendix K. Robust test of regression models (Chapter 5.2)

Table Appendix-9. Robust test with bootstrapping with 2000 samples (Chapter 5.2)

	B	*SE.B*	*p*	*95% CI*	*95% CI (bootstrap)*
Effect of decision aids on system cognitive effort (Step 2 in Table 5-7)					
Constant	3.65	0.63	0.00	(2.42, 4.89)	(2.43, 4.86)
Age	0.00	0.03	0.88	(-0.06, 0.05)	(-0.047, 0.04)
Gender	0.16	0.20	0.42	(-0.23, 0.54)	(-0.22, 0.52)
Education	0.16	0.18	0.37	(-0.2, 0.53)	(-0.17, 0.55)
PDTK	0.01	0.08	0.85	(-0.14, 0.17)	(-0.12, 0.18)
TAGS	**-0.81**	**0.22**	**0.00**	**(-1.23, -0.38)**	**(-1.23, -0.38)**
TAXO	**-1.78**	**0.22**	**0.00**	**(-2.2, -1.35)**	**(-2.19, -1.35)**
Effect of decision aids on selection accuracy (Step 4 in Table 5-7)					
Constant	6.94	1.19	0.00	(4.6, 9.29)	(4.67, 9.18)
Age	0.01	0.05	0.81	(-0.09, 0.11)	(-0.07, 0.1)
Gender	-0.02	0.37	0.96	(-0.76, 0.71)	(-0.81, 0.78)
Education	-0.20	0.35	0.56	(-0.89, 0.49)	(-0.81, 0.49)
PDTK	-0.09	0.15	0.52	(-0.38, 0.19)	(-0.39, 0.15)
TAGS	**0.75**	**0.41**	**0.07**	**(-0.06, 1.56)**	**(-0.04, 1.54)**
TAXO	**5.11**	**0.41**	**0.00**	**(4.3, 5.92)**	**(4.29, 5.91)**
Effect of decision aids and system cognitive effort on selection accuracy (Step 5 in Table 5-7)					
Constant	8.28	1.27	0.00	(5.77, 10.8)	(5.46, 10.53)
Age	0.01	0.05	0.83	(-0.09, 0.11)	(-0.075, 0.12)
Gender	0.04	0.37	0.92	(-0.69, 0.76)	(-0.75, 0.75)
Education	-0.14	0.34	0.68	(-0.82, 0.54)	(-0.77, 0.53)
PDTK	-0.09	0.14	0.54	(-0.37, 0.2)	(-0.34, 0.17)
TAGS	0.45	0.42	0.28	(-0.38, 1.28)	(-0.35, 1.28)
TAXO	**4.46**	**0.47**	**0.00**	**(3.53, 5.39)**	**(3.43, 5.54)**
SCE	**-0.37**	**0.14**	**0.01**	**(-0.64, -0.1)**	**(-0.67, -0.082)**
Moderating effect of rational decision-making style on the relations between decision aid taxonomy and selection accuracy (Step 2 in Table 5-9)					
Constant	3.40	1.93	0.08	(-0.41, 7.21)	(-0.60, 6.44)
Age	0.03	0.05	0.60	(-0.07, 0.12)	(-0.06, 0.11)
Gender	-0.03	0.37	0.93	(-0.76, 0.7)	(-0.81, 0.74)
Education	-0.30	0.35	0.40	(-0.98, 0.39)	(-0.95, 0.37)
PDTK	-0.09	0.15	0.53	(-0.38, 0.2)	(-0.37, 0.17)
TAGS	1.42	2.59	0.59	(-3.7, 6.53)	(-2.98, 6.79)
TAXO	**9.40**	**2.42**	**0.00**	**(4.63, 14.17)**	**(4.57, 14.20)**
RDS	**0.62**	**0.27**	**0.02**	**(0.09, 1.16)**	**(0.19, 1.20)**
TAGS x RDS	-0.13	0.47	0.78	(-1.06, 0.8)	(-1.09, 0.67)
TAXO x RDS	**-0.79**	**0.43**	**0.07**	**(-1.65, 0.06)**	**(-1.65, 0.053)**
Notes: PDTK: perceived design technique knowledge, TAXO: decision aid taxonomy, TAGS: decision aid tags, SCE: system cognitive effort, RDS: rational decision-making style.					

Appendix L. Tasks and interview questions used in the usability evaluation for ServiceDesignKIT 2.0

Task description

Please imagine you were a developer in a grocery company. Recently, your company decided to build an online shop. You were assigned to conduct user research in a short-term period. Please use servicedesignkit.org to search design techniques that can help you accomplish this task.

1. Please use the taxonomy and tag cloud to select appropriate design techniques and add them to your favorite list.
2. Use the technique advisor to get suggestion of design techniques and save them in the suggestion history. Review the suggestion history.
3. Submit a design technique with the following information.

   Technique name: Personas
   Description: Description
   Short Instruction: Instruction
   Reference: Reference
   Add a tag: Tag1
   Design Phase: Planning
   Time Dependency: Retrospective
   Duration: Short-Term Study
   User Participation: User Absence
   Evaluation Type: Group Discussion

Interview questions

Starting Question

1. How do you think of servicedesignkit.org? For example, is it helpful when you were searching design techniques?

Classification

2. Did the classification help you find the desired design techniques? (If not, why)
3. Did you have any problems in using the filtering?
4. In your opinion, how should the classification look like?

Technique advisor

5. Did the technique advisor help you find the desired design techniques? (If not, why)
6. Were the questions asked by the advisor understandable?
7. From your point of view, what kind of functions should the technique advisor also provide?

Individual feature

8. How do you think about the favorite list and advisory history?
9. How do you think about the user interface of suggesting new design techniques?
10. Are there other functions you would like to add to in individual model?

Intention to use

11. Will you use this platform in the future if you have to find design techniques for specific design situations?

Appendix M. Screenshot for the experiment training session (Chapter 6.2)

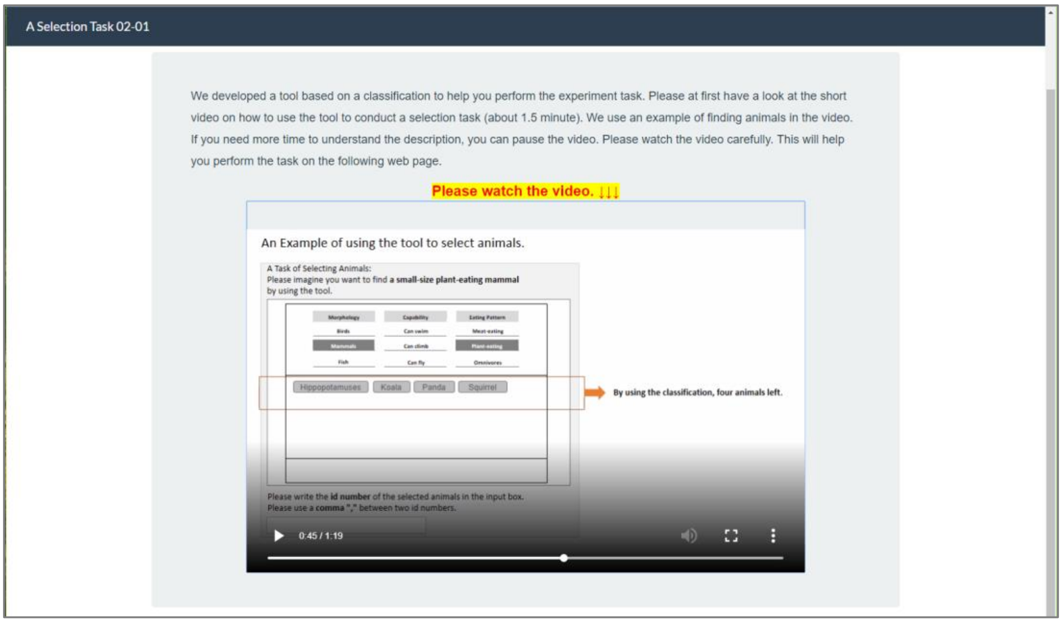


Figure Appendix-6. Screenshot of the description of using the taxonomy UI

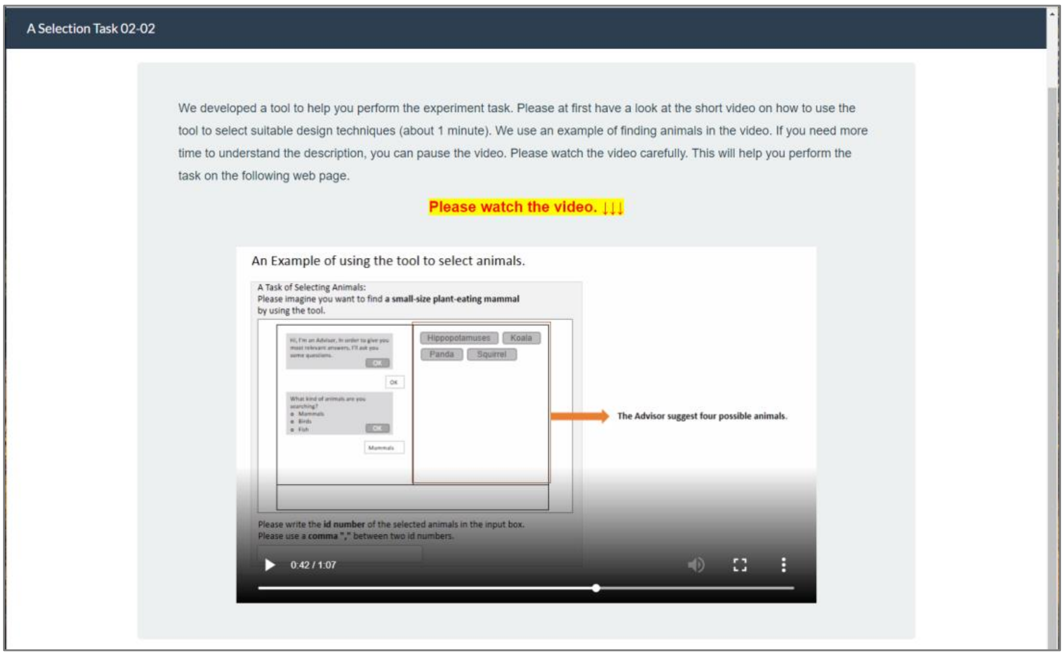


Figure Appendix-7. Screenshot of the description of using the natural language UI

Appendix N. Measurements and items used for evaluating the effect of taxonomy UI and natural language UI on the selection

Table Appendix-10. Measurements and items used in Chapter 6.2[1]

Measurement	Item
Rational Decision-making Style (Hamilton et al. 2016)	1. I prefer to gather all the necessary information before coming to a decision. 2. I thoroughly evaluate decision alternatives before making a final choice. 3. In decision making, I take time to contemplate the pros/cons or risks/benefits of a situation. 4. Investigating the facts is an important part of my decision-making process. 5. I weigh a number of different factors when making decisions.
Perceived Complexity (Thompson et al. 1994)	1. Using the [classification of design techniques/design technique advisor] takes too much time from my normal duties. 2. Working with the [classification of design techniques/design technique advisor] is so complicated, it is difficult to understand what is going on. 3. Using the [classification of design techniques/design technique advisor] involves too much time to click through all the categories. 4. It takes too long to learn how to use the [classification of design techniques/design technique advisor] to make it worth the effort.
Perceived Ease of Use (Wixom and Todd 2005)	1. The [classification of design techniques/design technique advisor] is easy to use. 2. It is easy to get the [classification of design techniques/design technique advisor] to do what I want it to do. 3. The [classification of design techniques/design technique advisor] is easy to operate.
Perceived Task Performance (Kim et al. 2009)	1. Using the [classification of design techniques/design technique advisor] improved my performance in the tasks. 2. Using the [classification of design techniques/design technique advisor] increased my productivity in the tasks. 3. Using the [classification of design techniques/design technique advisor] enhanced my effectiveness in the tasks. 4. Overall, using the [classification of design techniques/design technique advisor] is useful in selecting design techniques.

[1] In order to make the questions easy to understand for experiment participants, in the self-evaluation questionnaire, the term taxonomy and natural language dialog was replaced by classification of design techniques and design technique advisor.

Appendix O. Rotated factor loading (Chapter 6.2)

Table Appendix-11. Rotated factor loading (Chapter 6.2)

Item	RC1	RC2	RC3	RC4	h2	u2
Perceived Task Performance_01	**0.78**	0.1	-0.23	0.24	0.74	0.26
Perceived Task Performance_02	**0.8**	0.16	-0.16	0.22	0.73	0.27
Perceived Task Performance_03	**0.84**	0.15	-0.14	0.18	0.79	0.21
Perceived Task Performance_04	**0.86**	0.09	-0.19	0.15	0.81	0.19
Rational Decision-making Style_01	-0.04	**0.75**	0	0.11	0.58	0.42
Rational Decision-making Style _02	0.08	**0.85**	0.05	0.14	0.75	0.25
Rational Decision-making Style _03	0.12	**0.85**	-0.15	0	0.77	0.23
Rational Decision-making Style _04	0.31	**0.78**	-0.2	-0.06	0.75	0.25
Rational Decision-making Style _05	0.13	**0.67**	-0.07	0.2	0.51	0.49
Perceived Complexity_01	-0.35	-0.04	**0.75**	-0.1	0.7	0.3
Perceived Complexity_02	-0.17	-0.11	**0.83**	-0.24	0.78	0.22
Perceived Complexity_03	-0.2	-0.02	**0.79**	-0.21	0.71	0.29
Perceived Complexity_04	-0.04	-0.13	**0.86**	-0.21	0.8	0.2
Perceived Ease of Use_01	0.19	0.18	-0.34	**0.73**	0.72	0.28
Perceived Ease of Use_02	0.4	0.09	-0.17	**0.73**	0.74	0.26
Perceived Ease of Use_03	0.27	0.14	-0.35	**0.74**	0.77	0.23

Notes: Extraction methods: principal component factoring; rotation method: varimax rotation; h2: share of variance explained by factors; u2: share of variance not explained by factors; com: an index for the “complexity” of an item (not covered in this class).

Appendix P. Robust test of regression models (Chapter 6.2)

Table Appendix-12. Robust test with bootstrapping with 2000 samples (Chapter 6.2)

	B	*SE.B*	*p*	*95% CI*	*95% CI (bootstrap)*
H4a. Rational decision-making style has a moderating effect on the relationship between using the two UIs and perceived complexity.					
Constant	9.30	1.61	0.00	(-7.41, 0.52)	*(-7.22, 0.54)*
Age	0.01	0.03	0.59	(-0.05, 0.1)	*(-0.047, 0.097)*
Gender	0.07	0.13	0.59	(-0.29, 0.37)	*(-0.26, 0.37)*
Edu	-0.07	0.13	0.61	(-0.26, 0.4)	*(-0.22, 0.50)*
UI	**-3.37**	**0.92**	**0.00**	**(1.75, 6.27)**	***(1.44, 6.53)***
RDS	**-0.70**	**0.28**	**0.01**	**(0.22, 1.6)**	***(0.18, 1.57)***
UI X RDS	**0.62**	**0.17**	**0.00**	**(-1.15, -0.33)**	***(-1.18, -0.29)***
H4b. Rational decision-making style has a moderating effect on the relationship between using the two UIs and perceived ease of use.					
Constant	9.30	1.61	0.00	(6.12, 12.48)	*(6.60, 12.37)*
Age	0.02	0.03	0.59	(-0.04, 0.07)	*(-0.030, 0.069)*
Gender	0.07	0.13	0.59	(-0.19, 0.34)	*(-0.17, 0.30)*
Edu	-0.07	0.13	0.61	(-0.33, 0.2)	*(-0.41, 0.19)*
UI	**-3.37**	**0.92**	**0.00**	**(-5.18, -1.55)**	***(-5.14, -1.51)***
RDS	**-0.70**	**0.28**	**0.01**	**(-1.25, -0.14)**	***(-1.21, -0.20)***
UI X RDS	**0.62**	**0.17**	**0.00**	**(0.29, 0.95)**	***(0.31, 0.94)***
H5a. When using the taxonomy and natural language dialog UIs to select design techniques, perceived complexity has a negative effect on perceived task performance.					
Constant	5.51	0.71	0.00	(4.11, 6.9)	*(4.18, 6.69)*
Age	0.05	0.03	0.15	(-0.02, 0.11)	*(-0.01, 0.10)*
Gender	0.04	0.15	0.81	(-0.26, 0.34)	*(-0.32, 0.33)*
Edu	-0.07	0.15	0.65	(-0.37, 0.23)	*(-0.41, 0.24)*
PC	**-0.46**	**0.07**	**0.00**	**(-0.61, -0.32)**	***(-0.65, -0.26)***
H5b. When using the taxonomy and natural language dialog UIs to select design techniques, perceived ease of use has a positive effect on perceived task performance.					
Constant	1.00	0.78	0.20	(-0.54, 2.54)	*(-0.85, 2.70)*
Age	0.02	0.03	0.48	(-0.04, 0.08)	*(-0.04, 0.08)*
Gender	-0.01	0.14	0.92	(-0.29, 0.26)	*(-0.32, 0.28)*
Edu	-0.06	0.14	0.68	(-0.34, 0.22)	*(-0.30, 0.24)*
PEU	**0.70**	**0.08**	**0.00**	**(0.54, 0.87)**	***(0.53, 0.88)***
Notes: RDS: rational decision-making style; PC: perceived complexity; PEU: perceived ease of use					

List of Publications Included in this Thesis

Publications

Liu, X., He, S., and Maedche, Al. 2019. "Designing an AI-Based Advisory Platform for Design Techniques," in *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm, Sweden, pp. 1–16.

Liu, X., Werder, K., and Zhao, Q. (2018). "The Role of Cultural Differences when Using Different Classifications: An Experiment Design for Design Technique Selection," in *Proceedings of the 39th International Conference on Information Systems (ICIS),* San Francisco pp. 1–9.

Liu, X. (2018). "Selection Support of Digital Service Design Techniques for Design Novices," in *Proceedings of the Doctoral Consortium Papers Presented at the 30th International Conference on Advanced Information Systems Engineering (CAiSE),* Tallinn, Estonia, pp. 1–9.

Liu, X., Leung, E. T.-M., Toreini, P., and Maedche, A. (2018). "ServiceDesignKIT: A Web Platform of Digital Service Design Techniques," in *Proceedings of 13th International Conference on Design Science Research in Information Systems and Technology (DESRIST),* S. Chatterjee, K. Dutta, and R. Sundarraj (eds.), Cham: Springer, pp. 34–48.

Liu, X., Werder, K., and Maedche, A. (2016). "A Taxonomy of Digital Service Design Techniques," in *Proceedings of the 37th International Conference on Information Systems (ICIS),* Dublin, pp. 1–12.

Working Papers

Liu, X., Werder, K., and Maedche, A. (2019). A Taxonomy of Design Techniques for Digital Services. *Working Paper*.

Liu, X., Werder, K., and Maedche, A. (2019). Supporting Novice Designers' Decision-Making with Interactive Decision Aids – A Comparison of high and low Structured Decision Aids. *Working Paper*.

Liu, X., Rietz, T., and Maedche, A. (2019). Influence of UIs on the Novices' Selection Performance – A Comparison of Taxonomy and Natural Language Dialog. *Working Paper*.

Eidesstattliche Versicherung

gemäß § 6 Abs. 1 Ziff. 4 der Promotionsordnung des Karlsruher Instituts für Technologie für die Fakultät für Wirtschaftswissenschaften

1. Bei der eingereichten Dissertation zu dem Thema *Designing Decision Aids for Digital Service Design Technique Selection* handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich *bislang nicht* an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Karlsruhe, den *13.08.2019*

____

Xuanhui Liu