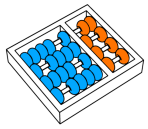
University of Campinas
Institute of Computing



Flavio Nicastro

A semiotic-informed approach to interface guidelines for mobile applications: a case study on Phenology data acquisition

Uma abordagem informada por semiótica para diretrizes de interfaces para aplicativos de dispositivos móveis: um estudo de caso na aquisição de dados fenológicos

CAMPINAS
2015

University of Campinas
Institute of Computing

Flavio Nicastro

A semiotic-informed approach to interface guidelines for mobile applications: a case study on Phenology data acquisition

Uma abordagem informada por semiótica para diretrizes de interfaces para aplicativos de dispositivos móveis: um estudo de caso na aquisição de dados fenológicos

Dissertation presented to the Institute of Computing of the University of Campinas in partial fulfillment of the requerements for the degree of Master in Computer Science.

Dissertação apresentada ao Instituto de Computação da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Supervisor/Orientador: Prof. Dr. Ricardo da S. Torres
Co-supervisor/Coorientadora: Prof. Dra. M. Cecilia C. Baranauskas

Este exemplar corresponde à versão final da Dissertação defendida por Flavio Nicastro e orientada pelo Prof. Dr. Ricardo da S. Torres.

Ricardo Torres

Supervisor

CAMPINAS
2015

Agência de fomento: Não se aplica
Nº processo: Não se aplica

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

N513s

Nicastro, Flavio, 1972-
A semiotic-informed approach to interface guidelines for mobile applications : a case study on phenology data acquisition / Flavio Nicastro. – Campinas, SP : [s.n.], 2015.

Orientador: Ricardo da Silva Torres.
Coorientador: Maria Cecilia Calani Baranauskas.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Computação.

1. Interação humano-computador. 2. Semiótica organizacional. 3. Fenologia. 4. Interfaces em aplicações para dispositivos móveis - Manuais, guias, etc. I. Torres, Ricardo da Silva,1977-. II. Baranauskas, Maria Cecilia Calani,1954-. III. Universidade Estadual de Campinas. Instituto de Computação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Uma abordagem informada por semiótica para diretrizes de interfaces para aplicativos de dispositivos móveis : um estudo de caso na aquisição de dados fenológicos
Palavras-chave em inglês:
Human-computer interaction
Organizational semiotics
Phenology
Mobile application interface - Handbooks, manual, etc.
Área de concentração: Ciência da Computação
Titulação: Mestre em Ciência da Computação
Banca examinadora:
Ricardo da Silva Torres [Orientador]
Debora Pignatari Drucker
Heiko Horst Hornung
Data de defesa: 10-04-2015
Programa de Pós-Graduação: Ciência da Computação

T E R M O D E A P R O V A Ç Ã O

Defesa de Dissertação de Mestrado em Ciência da Computação, apresentada pelo(a) Mestrando(a) **Flavio Nicastro**, aprovado(a) em **10 de abril de 2015**, pela Banca examinadora composta pelos Professores(as) Doutores(as):

Prof(a). Dr(a). Debora Pignatari Drucker
Titular

Prof(a). Dr(a). Heiko Horst Hornung
Titular

Prof(a). Dr(a). Ricardo da Silva Torres
Presidente

A Ata da Defesa, onde constam as assinaturas dos membros da banca, encontra-se no processo de vida acadêmica do aluno.

Dedicatória

Dedico este trabalho aos meus pais, minha esposa e aos meus filhos. Sem eles seria impossível ter chegado até aqui.

If you can't explain it simply, you don't understand it well enough.
Everything should be made as simple as possible, but not simpler.
(Albert Einstein)

Agradecimentos

This research was supported by the São Paulo Research Foundation (FAPESP) - Microsoft Research Virtual Institute (grants #2010/52113-5 and #2013/50155-0). Also, we have been benefited from funds of CNPq, CAPES, and FAPESP.

We would like also to thank the research group led by Profa. Cecilia Baranauskas, with special thanks to experts for their help and participation in the evaluation process conducted in this work.

Resumo

Dispositivos móveis são utilizados para aquisição de dados em diferentes domínios, como por exemplo, logística e aquisição de dados censitários. No entanto, a sua adoção em larga escala depende do desenvolvimento de aplicações com um cuidadoso *design* de interação. Nesta dissertação, revisitamos estratégias de *design* e avaliação de aplicações para dispositivos móveis e usamos a escada semiótica, artefato da Semiótica Organizacional, para organizar estas diretrizes. Propomos um conjunto de *guidelines* com questões, orientados pela semiótica, para a avaliação de interfaces de aplicações em dispositivos móveis. Também propomos uma metodologia para avaliação de interfaces de aplicações em dispositivos móveis baseada no conjunto de *guidelines* proposto. Demonstramos a utilização deste conjunto de guidelines na avaliação de quatro aplicações desenhadas para a aquisição em campo de dados fenológicos com o uso de dispositivos móveis.

Abstract

Portable devices have been experimented for data acquisition in different domains, e.g., logistics and census data acquisition. Nevertheless, their large-scale adoption depends on the development of effective applications with a careful interaction design. In this dissertation, we revisit existing strategies for mobile application design and evaluation and use the Semiotic Ladder from Organizational Semiotics as an artifact to organize a set of guidelines. We propose a set of semiotic-informed guidelines with questions for evaluation of mobile application interfaces. We also propose a methodology for the evaluation of mobile application interfaces based on the proposed guidelines set. We demonstrate the use of the proposed methodology in the evaluation of four mobile application interfaces designed for phenological data acquisition in the field.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Portable devices have been adopted in different domains to support data acquisition [1,14, 42]. Key motivations for their use rely on the associated low costs, recent improvements in the hardware robustness, the incorporation of different sensors (e.g., for location, audio, image, and video acquisition) that provide useful contextual information for different purposes, and the availability of easy-to-use frameworks for developing applications.

However, the effective adoption of portable devices depends on the use of applications with careful interface design. An appropriate design should be associated with reduced mental and physical stress, reduced learning curve, and improved device operability [12]. The design and implementation of such interfaces deal with many constraints: small-size screens, data entry models, connectivity issues, and limited resources. Moreover, there are other factors underlying the user-system interaction that should also be considered, such as the social implications of changing work practices. In this sense, the definition and use of appropriate interface design guidelines may help application developers to address part of these challenges.

Information and Communication Technology has evolved rapidly, shaping our relationships in the world (e.g., economic, social, laboral, interpersonal, and ethical). The relationship between people and information is changing at the same pace, mediated by that technology. Within this scenario, and as part of it, designing or evaluating an application demands a systemic view on the prospective product of that technology. For this systemic view on the design of computer-based applications, Organizational Semiotics (OS) [18] has been a fundamental theoretical frame of reference for our work.

By making use of several design cues, indicators, and signs, Semiotics, the doctrine of signs, enables us to search for a more accurate understanding of information as properties of signs. Anything that stands for something or is used to mean something to someone is an example of a sign: words, sentences, traffic lights, diagrams, a wave, a facial expression. Adopting Baranauskas' perspective to design [4, 5], we take Semiotics beyond the study of how we use signs to communicate, to include shared knowledge and mutual commitment that establishes communication in the design process. In this sense, information, understood as signs, could be operated in distinct levels, meaning different operations a person can do upon the sign. These levels are represented as steps of a Semiotic Ladder (SL), or views of a semiotic framework [38].

In this dissertation, we revisit existing mobile application design and evaluation strate-

gies and use the Semiotic Ladder from OS as an artifact to organize guidelines to support the design and the evaluation of mobile application interfaces. To the best of our knowledge, this is the first attempt to categorize existing guidelines according to the SL. The final guidelines set proposed, composed of 27 guidelines, is expected to support mobile interface designers, developers, and evaluators in their daily tasks. We also propose a method for mobile application interface evaluation based on the defined set of guidelines. Another contribution of this work is that the proposed method can be conducted with the most interested parties: people from the application domain, not necessarily with knowledge on semiotics or on interface evaluation procedures.

We demonstrate the use of the proposed methodology in the evaluation of four mobile application interfaces recently proposed in the e-Science domain. In these applications, we are interested in supporting data acquisition upon plant phenology in the field. Plant phenology concerns the study of recurrent life cycle events and its relationship to climate [36]. This discipline has been recognized as a strategic approach to climate change research [36]. Plant phenology studies are based on a well-defined methodology that has as main objective the identification and understanding of temporal changes in reproductive or vegetative events [22]. Plant phenology studies depends, therefore, on the continuous acquisition and analysis of data over time. The common approach is the direct observation of plant individuals in the field at regular intervals (e.g., monthly or weekly) and the identification of phenophases (e.g., leafing, budding, flowering, and ripening) [22]. One widely adopted approach for data acquisition relies on using a qualitative method to assess the presence or absence of phenophase or using a quantitative method that assigns a different number (usually 0, 1, or 2) to a phenophase, depending on its intensity [25]. Usually, phenophase intensities are registered on paper sheets (in the field) and later inserted into digital spreadsheets (in the laboratory). This task is time consuming and error prone. These issues have motivated the investigation upon the use of portable devices to support the phenological data acquisition process.

The remaining of this text is organized as follows. Chapter 2 provides the background on related topics. Chapter 3 introduces the proposed semiotic-informed guidelines set to support the evaluation of mobile application interfaces, while Chapter 4 describes a case study concerning the use of the proposed guidelines in the evaluation of application interfaces proposed for phenological data acquisition. Finally, Chapter 5 presents our conclusions and directions for future work.

Chapter 2

Background and Literature Review

2.1 Phenology

Several research initiatives have recognized the importance of studying environmental changes. This is for example the main objective of Phenology, which investigates natural recurring phenomena and its relation to climate [36]. Phenology studies are dedicated to the observation of living beings and their relationship with meteorological data [29]. In the context of plant phenology, the budding of the leaves and the senescence are examples of important stages in the cycles of plants that usually are monitored in phenology studies dedicated to understanding several ecosystem processes such as growth, water status, gas exchange, and nutrient cycling [24,31]. In another scenario, plant phenology is important to the estimation of carbon balance and land productivity [16, 19, 34].

One important research venue in the Phenology is concerned with the specification and implementation of novel technologies for phenological observation [3,26,32,33,39,40]. This is, for example, the main objective of the e-phenology project,[1] which is a pioneer project in Brazil dedicated to performing phenology studies by considering information obtained from vegetation images. It is a multidisciplinary project that combines research in Computer Science and Phenology. The project is developed in the context of a collaboration involving the Laboratory of Phenology, Institute of Biosciences (IB), São Paulo State University "Júlio de Mesquita Filho" (UNESP) and the Recod Laboratory (Reasoning for Complex Data), in the Institute of Computing, University of Campinas (Unicamp). The main objectives are: (1) the use of the new technologies of environmental monitoring, (2) the creation of a protocol to a program of long-term phenological monitoring in Brazil, and (3) the proposal of models, methods, and algorithms to support the management, integration, and analyses of phenological data.

The current studies of the e-phenology project are based on a Cerrado area, located in the region of Itirapina, São Paulo. The phenological data in this area have been obtained by biologists from Phenology Laboratory of UNESP since September 2004, and are related to the observations of the occurrence of biological phenomena in the life cycle of plants, called phenophases.

The observed phenological phases are: flowering, fruiting, budding, and leaf fall. The

[1] `http://www.recod.ic.unicamp.br/ephenology/` (As of March 2015).

production of flowers is divided into buttons and flowering itself, or anthesis (flower opening); the production of fruit is divided into development of fruits and unripe fruit, and the fruitification into period of ripe fruit and "ready to spread." Every month researchers perform the field observation of several individuals. Each observer takes a clipboard with printed sheets similar to that presented in Appendix C. On these sheets, they register information of different individuals such as their identification number, family, type, and location. It also contains fields for the annotation of the intensity of the observed pehnophases of each individual, represented by the letters: Bot – Button, Ante – anthesis, FV – green fruit, FM – ripe fruit, Brot – budding and Qued – leaf fall. The stages are quantified according to three intensity classes: 0 – indicates that the phenology is not present; 1 – indicates presence of phenology at a lower intensity; 2 – indicates the presence of phenology at a higher intensity. After the researchers complete the observations of all individuals, they register observed data into digital spreadsheets. Later these worksheets are analyzed in order to detect inconsistencies in the data collected. For example, it is not correct the identification of flower without having registered the phenophase button in a previous observation. If an inconsistency is detected, it is considered that there is a problem of observation and based on the current observation, the value of the previous observation is modified. This process of collection began in 2004 and today, phenology studies have been performed by analyzing about a million records related to more than 2000 plant individuals. Recently, a database model was proposed to store all data managed in the context of the e-phenology project [20].

2.2 Mobile Interfaces

New technologies provide power to people who are able to handle them properly. A technology becomes widely accepted and effectively used when well designed, and this means meeting the needs and capabilities of a target group. A key factor for the success of a particular technology relies on its interface with potential users. In this particular matter, one important concern consists in defining appropriate mechanisms to improve the way how people can use technology to think and communicate, observe and decide, calculate and simulate, and discuss and design. The Human-Computer Interaction (HCI) area addresses these issues by proposing appropriate approaches for the design of interface that may help people so that they can perform their activities with productivity, safety and satisfaction [6].

With the exponential use of mobile devices, HCI researchers have dedicated to understanding and develop standards that improve the usability and the quality of interaction between human and portable devices. In particular, in this dissertation, we aim to define an appropriate set of guidelines for mobile application interface design and evaluation based on the identification of existing research in the area. Initially, existing guidelines were selected from publications associated with ACM and IEEE conferences and journals. At this stage about 50 papers addressing mobile HCI were identified as relevant. After a read over them, a set of 24 papers were selected for a deep analysis. Finally, a set of 12 articles were chosen as those most important for this work.

Table 2.1 summarizes some relevant work in the literature, without exhausting the subject. The literature studies point to different aspects that are relevant in evaluating mobile application user interfaces. In this work, we select and refine guidelines and questions proposed in those studies, classifying them according to the different semiotic layers.

Table 2.1: Literature overview on mobile interface design.

References	Overview
Nayeb et al., 2012 [23]	- this work analyses the state-of-the-art concerning the evaluation of the usability of mobile applications; - presents a methodology for usability evaluation; - suggests that there is little scientific research in this area.
Radia et al., 2012 [28]	- presents guidelines based on the latest research in industry and academia; - search design and development of successful mobile applications that can utilize the capabilities of next generation cellular network; - presents a model for developing client-server applications based on 4G technologies.
Zamzami and Mahmud, 2012 [41]	- states that there is little research focused on assessing the information quality on smartphone interfaces; - examines three main areas: mobile interface design, information quality, and user satisfaction.
Rauch, 2011 [30]	- discusses differences in usability research focused on desktops compared with what has been done for mobile devices; - summarizes emerging trends in usability studies for mobile devices; - suggests best practices for converting documentation to Kindle-compatible .MOBI format.
Ayob et al., 2009 [8]	- proposes a three-layer design model for mobile applications based on four existing guidelines.
Hussain and Kutar, 2009 [15]	- presents a usability metric framework for mobile phone applications; - proposes 6 guidelines with 21 questions and 30 metrics for usability evaluation; - include quality characteristics from ISO 9241-11.
Ryu, 2005 [35]	- presents usability questionnaires for electronic mobile products and decision making methods; - proposes an evaluation questionnaire containing 72 items.
Gong and Tarasewich, 2004 [13]	- presents guidelines for handheld mobile device interface design; - is based on golden rules of interface design. - one of the first attempt to organize a set of guidelines for the design of mobile devices interfaces.

In the following, these works are briefly outlined. In the paper of Nayeb et al. [23], the authors present the state of the art concerning the evaluation of the usability of mobile

applications. They point out that is always important to consider three aspects of usability for all types of software: a) More efficient to use: take less time to complete a particular task; b) Easier to learn: operations can be learned observing the object, and c) More user satisfaction: meets user expectations. Also, they attest that referring to evaluation methodology, three types are currently used in studies of mobile usability: 1) Laboratory experiments; 2) Field studies, and 3) Hands-on measurements. They state that there is little scientific research in this area and an evaluation methodology is presented to fill this gap.

In the paper of Radia et al. [28], the authors present some guidelines based on recent research from academia and industry, seeking the design and development of mobile applications that can use the capabilities of the next generation of mobile network. They present a model for application development using a 4G-based client-server. Referring to the guidelines, the most important contribution from this paper to our work, the authors discuss about the categorization of guidelines into three broad classes. The first class, "General UI Guidelines for Mobile Applications" is concerned with user interfaces. Some guidelines proposed in this class include: 1) provide shortcuts for experienced users and wizard for new users; 2) allow to maintain control by having the ability to control the application (or abort it) at any point; 3) create good dialogues by creating predictable and intuitive sequences of interaction with the application; 4) minimize dependence on user's memory through grouping information in "chunks" at a time, limiting the need for scrolling; among others. The second class, "Mobility Guidelines" is responsible for specific guidelines when dealing with mobility. Some examples: 1) allowing for multimodal interactions with the device; 2) allowing for convenient use with the ability to handle multiple and frequent interruptions with limited attention from the user; 3) providing an ability to synchronize the application with desktop and cloud data stores; 4) allowing privacy for single or multiple users; among others. And the third class, "Organizational Guidelines", is concerned with enterprise-specific and corporation guidelines. Some examples in this class are: 1) consistency with the organization's standards and systems; 2) support for business models; among others. The guidelines presented in the first and the second classes were very important for the definition of the set of guidelines proposed in this dissertation.

A quite similar view by Nayeb et al. [23] is presented by Zamzami and Mahmud [41]. In their work, the authors state that there is little research to evaluate the quality of information in smartphone interfaces. They perform their analysis considering three key areas: design mobile interfaces, information quality, and user satisfaction. Regarding to mobile user interface design, they attest that there are three principals that need to be followed in designing user interface: 1) let the users be in control of the interface; 2) reduce user's memory load; and 3) make the user interface consistent. Also, they discuss some concepts concerning user satisfaction.

Also showing that much has to be researched and developed in the HCI area for mobile devices, Marta Rauch [30] addresses the distance of desktops in usability research when compared to what has been done for mobile devices. She summarizes the emerging trends in usability for mobile devices and suggests "best practices" for converting documentation to Kindle-compatible *.MOBI* format. She also presents a study based on the development

of documentation on Kindle, Tablets, and Smartphones. Regarding to the guidelines, she brings the emerging usability guidelines for applications on these type of devices. Some usability guidelines pointed out include: 1) Consider the unique issues of mobile usability; 2) Analyze mobile user tasks; 3) Determine the target mobile device; 4) Help users avoid inadvertent actions; among others. She also discusses key requirements for professional communicators remember the guidelines. Example: for guideline 1: "When creating user assistance for mobile devices, ensure that the design is conducive to mobile use"; for guideline 2: "Determine whether a majority of your target audience will use your applications and documentation on mobile devices."

In Ayob et al. [8], the authors, in turn, propose a design model of three layers for mobile applications based on four existing guidelines. First, they raise the main issue of designing mobile application: "how to display all the information and elements in the small screen of mobile device?". The authors also present a discussion and a comparison about guidelines which is based on four existing set of guidelines named as: 1) Shneiderman's Golden Rules of Interface Design [37]; 2) Seven Usability Guidelines for Mobile Device; 3) Human-Centered Design (ISO Standard 13407), and 4) Mobile Web Best Practices 1.0 (W3C).[2] Some examples of guidelines from these four groups are: From group 1: 1) Enable frequent users to use shortcuts; 2) Reduce short-term memory load; 3) Design for small devices; among others. From group 2: 1) Meet user's need quickly; 2) Make user input as simple as possible; 3) Only show essential information; among others. From group 3: 1) Understand and specify the context of use; 2) Produce design and prototypes; among others. And for group 4: 1) User input; 2) Page layout and content; among others. Based on these comparisons, they propose a new guideline set divided into three layers called: 1) Analysis; 2) Designs and 3) Testing. Some examples of the guidelines proposed: From layer Analysis: 1) Identify and document user's task; 2) Define the use of the system; among others. From layer Design: 1) Enable frequent users to use shortcuts; 2) Design for multiple and dynamic contexts; among others. And from the third layer, Testing: 1) Usability testing; 2) Field studies; among others.

Other important work that addresses usability for mobile application is from Hussain and Kutar [15]. In their work, they present a usability metric framework for mobile phone application. This framework is composed of three main measures: Effectiveness, efficiency, and satisfaction. Each measure has two guidelines. For example, for Efficiency, the guidelines Time Taken and Features are defined. In total, they proposed six guidelines, based on which 21 questions are defined. From these questions, 23 metrics for usability evaluation were developed. An example of this relation between measures, guideline, question, and metric considers the following structure: Measure: Effectiveness; Guideline: Simplicity; Questions: Is it simple to key-in the data? Does the application provide virtual keypad? Is the output easy to use? How easy is it to install the application? Is the application easy to learn?; Metrics: Time taken to key-in the data; Provide/not provide virtual keypad for touch screen device; Provide/not provide help when necessary; Optimized/not optimized the screen size; - Rating scale for satisfaction on output; Time taken to install; The number of interaction occurred while installing the application; Successful/unsuccessful installation; Time taken to learn each task; Number of mistakes

[2] `http://www.w3.org/TR/mobile-bp/` (As of June 2015)

while learning. Also, the authors present a table with the most popular guidelines obtained from literature. Some of these guidelines are: 1) Completeness: the extent or completeness of user's solutions to tasks; 2) Less or no error: errors made by the user during the process of completing the task; 3) Simple: the application should be straightforward; 4) Ease to learn: the user interface must be designed for user to learn easily; among others.

In other similar study, Ryu [35] presents usability questionnaires for electronic mobile products and decision making methods. He also proposes an evaluation questionnaire containing 72 items for usability evaluation of mobile products. This questionnaire was divided into six groups: 1) Ease of Learning and Use; 2) Helpfulness and Problem Solving Capabilities; 3) Affective Aspect and Multimedia Properties; 4) Commands and Minimal Memory Load; 5) Control and Efficiency, and 6) Typical Task for Mobile Phone. One example of each group: 1) Is it easy to learn to operate this product? 2) Are the messages aimed at prevent you from making mistakes adequate? 3) Is this product attractive and pleasing? 4) Do the commands have distinctive meanings? 5) Is the data display sufficiently consistent? and for group 6) Is it sufficiently easy to operate keys with one hand? Some of these questions were used to compose our questionnaire for assessing the set of guidelines proposed in this dissertation.

Finally, Gong and Tarasewich [13] published in 2004 one of the first attempt to define a set of guidelines for the design of mobile device interfaces. They presented a guideline set for handheld mobile device interface design based on traditional guidelines for desktop user interfaces, the golden rules of interface design, and research with mobile device and applications. Some of the guidelines proposed in this study are: 1) Offer Informative Feedback; 2) Support Internal Locus of Control; 3) Design for multiple and dynamic contexts; 4) Design for limited and split attention; among others.

2.3 Semiotic Ladder

The Semiotic Ladder (SL) (illustrated in Figure 2.1) consists of six steps representing views on signs from the perspective of the physical world, empirics, syntactics, semantics, pragmatics, and the social world. The physical, empirics, and the social world are Stamper's [38] contribution upon the traditional semiotic approach. Stamper introduced it in his work, as a way of looking at meaning, communication, and information from a semiotic perspective [11].

The *Social World* is the layer in which we analyse the consequences of the use of signs in human activities. It deals for example with beliefs, expectations, commitments, law, and culture. *Pragmatics* is the layer studying the intentional use of signs and behavior of agents. Issues related to the intention and negotiation are objects of the pragmatic. *Semantics* deals with the relationship between a sign and what it refers to (its meaning); signs in all modes of signification. *Syntactics* deals with the combination of signs without considering their specific meaning. *Empirics* deals with the static properties of signs, when media and different physical devices are used. Finally, *Physical World* works with the physical aspects of signs and their marks (e.g., infrastructure issues).

In summary, the top three steps of the SL are related to the use of signs, how they

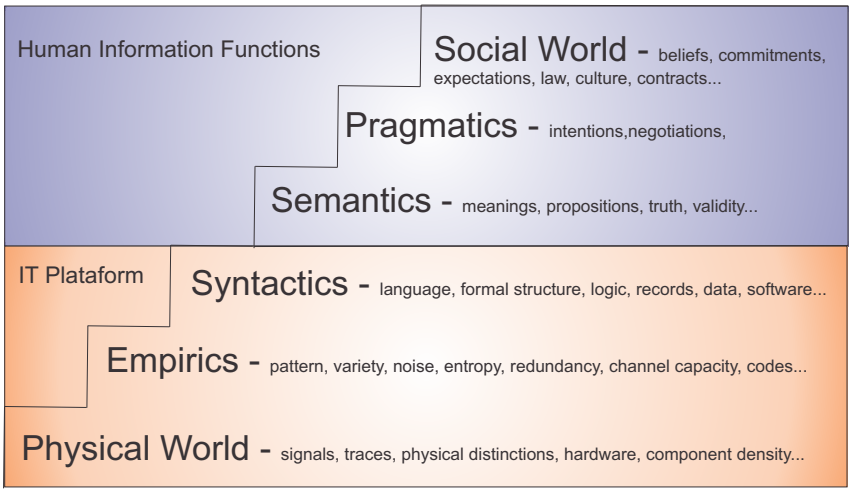


Figure 2.1: Semiotic Ladder steps. Figure adapted from Stamper [11].

work in communicating meanings and intentions, and the social consequences of their use. The three lower steps, in turn, answer questions related to how signs are structured and used, how they are organized and conveyed, and what physical properties they have, among others.

In the context of our study, the SL is an artifact that has been used for organizing the guidelines for system evaluation, covering aspects from its technological infrastructure (physical world, empirics, syntactic layer) to the system of human information (semantic layer, pragmatics, and social world). Therefore, this artifact supports both a wide and deep view of the different aspects that may be considered when evaluating interfaces for mobile devices. Other uses of the SL in different domains can be found in the literature, such as [27].

Chapter 3

Mobile Application Interface Evaluation

In this work, we propose a set of guidelines for supporting the evaluation of mobile application interfaces. We discuss the process used for defining this guidelines set in Section 3.1. Another contribution of this work refers to the proposal of an evaluation methodology that takes advantage of the proposed guidelines set. This methodology is described in Section 3.2.

3.1 Proposed Guideline Set

From the analysis of literature concerning the mobile application design and evaluation strategies based on publications associated with ACM and IEEE conferences and journals (see Table 2.1), we classified existing guidelines and questions according to the Semiotic Ladder. This step generated a total of 147 guidelines, distributed as follows: 15 into Physical World; 10 into Empirics; 58 into Syntactics; 26 into Semantics; 20 into Pragmatics, and 18 into Social World.

A novel set of guidelines was defined by merging similar ones and by discarding those considered non-pertinent to the design of mobile applications (13 in total). Examples of removed guidelines include questions such as:

- "Is it easy to change the ringer signal?",
- "Is it easy to check missed calls?",
- "Is it easy to check the last call?",
- "Is it easy to send and receive short messages using this product?",
- "Is it easy to use the phone book feature of this product?",
- "Can you personalize ringer signals with this product?".

Although those are general questions important to evaluate interaction and communication with the mobile device, they are not relevant for a mobile application interface

evaluation in the domain considered. The resulting set included 7 guidelines into the Physical World, 9 into Empirics; 16 into Syntactics, 10 into Semantics, 28 into Pragmatics, and 9 into the Social World, totalizing 79 guidelines and questions.

Finally, these guidelines and their classification were re-evaluated and refined by 10 experts in the fields of OS and HCI, to eliminate redundancies and to come with more appropriate descriptions and a more cohesive set, resulting in a set of 27 guidelines distributed as follows: 4 into the Physical World, 4 into Empirics, 7 into Syntactics, 3 into Semantics, 6 in Pragmatics, and 3 into the Social World. Figure 3.1 illustrates the whole process for refining and organizing the guidelines. The final proposed guidelines set is presented in Table 3.1.

Finally, a questionnaire was proposed to support the evaluation of mobile applications according to each defined guideline. The proposed questionnaire is presented in Tables 3.2, 3.3, 3.4, 3.5, 3.6, and 3.7. Each table refers to a different SL step. Note also that, more than one question can be used for a particular guideline in some cases. For example, for the fourth guideline related to the Physical World (*"Create design suitable for small devices with touch screen"*) – see Table 3.1, two questions are defined for the evaluation process: *"Are pictures on the screen of satisfactory quality and size?"*; *"Does the application provide appropriate menu button for touch screen?"*.

For each question, the evaluator should indicate, in a Likert scale (from 1 to 5), whether the application is attending to the guidelines recommendations. In the scale: 1 means that the evaluator strongly disagrees; 2, disagrees; 3, neither agrees nor disagrees; 4, agrees; 5, strongly agrees.

3.2 Evaluation Methodology

This section presents the proposed methodology for the evaluation of mobile application interfaces (Section 3.2.1) and provides an overview of created tools to support the use of the proposed methodology (Section 3.2.2).

3.2.1 Evaluation Methodology Description

The methodology is divided into four steps: definition of groups of evaluators; definition of application interfaces to be evaluated; evaluation of application interfaces; and performance of statistical tests.

Figure 3.2 presents the workflow of the proposed evaluation methodology. The first step (module labeled as A in the figure) is concerned with the definition of the application interfaces that will be evaluated. In our particular methodology, we consider the possibility of evaluating both running and *mockup* interface designs. The second step (module B) refers to the definition of groups of evaluators. In our methodology, we consider that evaluators with different background and expertise can be invited to take part in the evaluation process. For example, experts from a target domain (e.g., Phenologists)

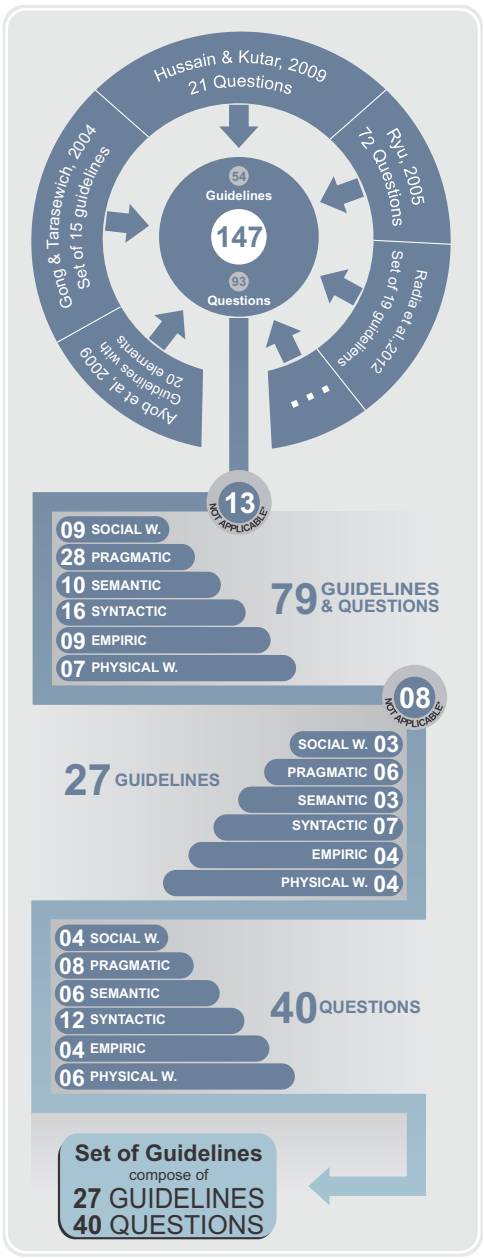


Figure 3.1: The guidelines refinement process. The resulting set, composed of 27 guidelines, is presented in Table 3.1, while the 40 questions proposed to evaluate mobile interfaces according to the defined guidelines are presented in Tables 3.2–3.7.

and experts of HCI could be invited to evaluate mobile application interfaces (e.g., two prototypes named 1 and 2) to be used for data acquisition in the field. In this scenario, two groups of evaluators are considered.

The third step (module C) is concerned with the actual evaluation. In this step, the evaluators use the defined application (result of step A) according to a pre-defined task. For example, evaluators use the application to register data collected regarding the intensity of specific phenophases of plants. After the evaluator performs the task in one or more application or prototypes, s/he fills out a questionnaire, based on the proposed guideline (see Section 3.1). The provided answers will be used to assess how certain

Table 3.1: Proposed Guidelines Set composed of 27 guidelines.

Step on the Semiotic Ladder	Guideline Description
Physical World	1 - Provide adequate contrast 2 - Provide methods for easy and functional data entry 3 - Create design suitable for small devices with touch screen 4 - Easy operation with one hand
Empiric	5 - Provide adequate response time and information display 6 - Keep recent data for reuse 7 - Facilitate data exchange with other applications 8 - Provide automatic application update
Syntactic	9 - Adequately provide information on system resources 10 - Give control to the user of the application 11 - Maintain consistency in the standards used both in data presentation as in how to perform each task 12 - Facilitate the navigation between screens and information 13 - Keep the user informed of what is happening through constant feedbacks 14 - Provide shortcuts and wizards 15 - Reduce mental efforts and memory requirements
Semantic	16 - Provide output of data easy to use 17 - Provide appropriate documentation by means of manuals and helps 18 - Design clear and understandable interfaces
Pragmatic	19 - Facilitate the discovery of new functionality 20 - Design the application thinking about simplicity 21 - Develop the application thinking in multiple contexts 22 - Provide feedback to aid the prevention of errors and troubleshooting as well as provide means for reversing actions 23 - Allow customization of the application by users 24 - Facilitate application learning
Social World	25 - Implement security and privacy controls 26 - Know your target audience in order to raise the application requirements, needs and “intrinsic” desires 27 - Implement controls to avoid risks while using the application in motion (driving, walking, etc.)

Table 3.2: Questionnarie Proposed - Physical World

Guideline	Question
1	1-Is the backlighting feature for the keyboard and screen appropriate in all contexts?
2	2-Does the application provide a virtual keypad? 3-Does the application provide voice assistance?
3	4-Are pictures on the screen of satisfactory quality and size? 5-Does the application provide appropriate menu button for touch screen?
4	6-Is it sufficiently easy to operate keys with one hand?

Table 3.3: Questionnarie Proposed - Empiric

Guideline	Question
5	7-Are the response time and information display fast enough?
6	8-Are data items kept short?
7	9-Are exchange and transmission of data between this product and other products (e.g., computer, PDA, and other mobile products) easy?
8	10-Does the application provide automatic update?

Table 3.4: Questionnarie Proposed - Syntactic

Guideline	Question
9	11-How much information about system resources was displayed?
10	12-Are the HOME and MENU buttons sufficiently easy to locate for all operations? 13-Are the letter codes for the menu selection designed carefully?
11	14-Are the color coding and data display compatible with familiar conventions? 15-Can all operations be carried out in a systematically similar way?
12	16-Is the organization of information on the application screen clear? 17-Is it easy to navigate between hierarchical menus, pages, and screen?
13	18-Does application provide feedback (haptic, audio, visual, etc.) constantly in order to keep the user engaged and attentive?
14	19-Does application provide shortcuts for experienced users and wizards for new users?
15	20-Does this application enable the quick, effective, and economical performance of tasks? 21-Does interacting with this application require a lot of mental effort? 22-Is it easy for you to remember how to perform tasks with this application?

interface is adherent to the items in the proposed guideline and how the interface supports the requirements of each step of the Semiotic Ladder. An example of one questionnaire is shown in Figure A.1.

In the fourth and last step (module D), statistical tests are performed. These tests have the objective of validating raised research questions. The first issue in this case refers to the definition of the research questions that will be addressed in the evaluation process. Examples include: "are there differences in answers among the group of evaluators?" (i.e., "is the evaluation of Prototype 1 by the group of HCI experts different from the evaluation conducted by Phenology experts?"). For answering these questions, a set of statistical tests has to be conducted based on the characteristics of the questions. Table 3.8 shows some commonly used significance tests for comparing means and their application context.

Table 3.5: Questionnarie Proposed - Semantic

Guideline	Question
16	23-Is the output data easy to use?
17	24-Are the documentation and manual for this application sufficiently informative? 25-Does the application provide appropriate help?
18	26-Is feedback on the completion of tasks clear? 27-Is the interface with this application clear and understandable? 28-Is the design of the graphic symbols, icons and labels on the icons sufficiently relevant?

Table 3.6: Questionnarie Proposed - Pragmatic

Guideline	Question
19	29-Is discovering new features sufficiently easy?
20	30-Is it easy to access the information that you need from the application?
21	31-Does application allow convenient use with the ability to handle multiple and frequent interruptions with limited attention from the user? 32-Does design of application is suitable for multiple contexts (home, business, travel, etc.) including support for runtime adaptation?
22	33-Are the error messages effective in assisting you to fix problems? 34-Are the messages aimed at prevent you from making mistakes adequate?
23	35-Does application provide the ability to personalize the application to suit the user?
24	36-Is it easy to learn to operate this application?

Table 3.7: Questionnarie Proposed - Social World

Guideline	Question
25	37-Does application allow privacy and security control for single or multiple users?
26	38-Does application have all the functions and capabilities you expect it to have? 39-Is this application attractive and pleasing?
27	40-Is the application secure to use while driving or walking?

In Table 3.9 we can see some possible questions with which each statistical test is used. For example, to answer the question "are there differences among the groups of evaluators?" a t test is the best choice.

The research questions considered in our methodology can be grouped into two cate-

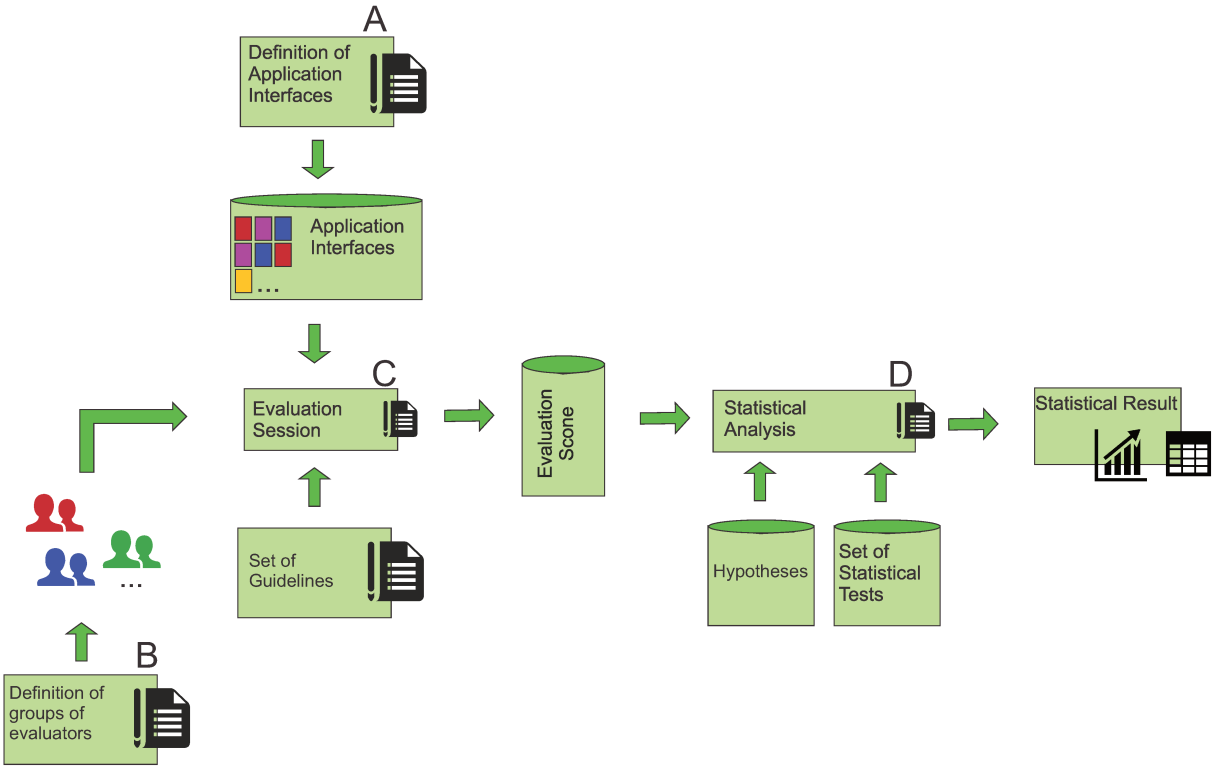


Figure 3.2: Workflow of the proposed methodology.

Table 3.8: Summary of statistical tests. Adapted from Lazar et. al [17].

Experiment Design	Independent Variables (IV)	Conditions for each IV	Types of test
Between group	1 1 2 or more	2 3 or more 2 or more	Independent Samples t test One-way ANOVA Factorial ANOVA
Within group	1 1 2 or more	2 3 or more 2 or more	Paired-samples t test Repeated Measures ANOVA Repeated Measures ANOVA
Between- and within-group	2 or more	2 or more	Split-plot ANOVA

Table 3.9: Examples of statistical tests for different types of questions.

Question	Type of test
Are there differences among the groups of evaluators	t test
Are there differences among the prototypes	Repeated measures ANOVA

gories: evaluator-centered and prototype-centered. The evaluator-centered questions aim to verify if there are differences among the groups of evaluators considering

1. their general evaluation of the prototypes;
2. their evaluation of a particular prototype;
3. their evaluation of prototypes using a particular Semiotic Ladder step; and

4. their evaluation of a particular prototype according to a specific Semiotic Ladder step.

The prototype-centered questions, in turn, aim to verify if there are differences among the prototypes considering

1. the complete set of guideline questions, i.e., all Semiotic Ladder steps altogether; and

2. a particular Semiotic Ladder step.

In Table 3.10, we show an example of results of t tests used to confirm the first question. In this example, both groups of evaluators have the same view of Prototype 1, but their evaluation differs with regard to Prototypes 2, 3, and 4.

Table 3.10: Examples of results of t tests: "Are there differences comparing the group of evaluators?"

Semiotic Ladder Step	All Prototypes	Prototype 1	Prototype 2	Prototype 3	Prototype 4
All Steps	Yes	No	Yes	Yes	Yes
Physical World	Yes	No	No	No	No
Empiric	Yes	No	No	No	Yes
Syntactic	Yes	No	No	Yes	Yes
Semantic	Yes	No	Yes	No	No
Pragmatic	Yes	No	No	Yes	Yes
Social World	Yes	No	No	No	No

3.2.2 Supporting Tools

We have created a set of tools to support the use of the evaluation methodology. Figure 3.3 presents the evaluation methodology workflow using the implemented tools. After the evaluators answer to the questionnaire, the scores are registered in a database, named *research database*. This database is linked to an R application, which is responsible for performing the statistical tests. In Figure B.1 (Appendix B) we show the Entity-Relationship Model of the research database. In this database, one query is defined to obtain the questionnaire results from the database. In the R application, scripts are executed using this query in order to obtain the answers for the key research questions. Figure B.2 (Appendix B) shows the script to create the research database, while Figure B.3 shows the implemented query. Figures B.4 and B.5, in turn, show the associated R scripts.

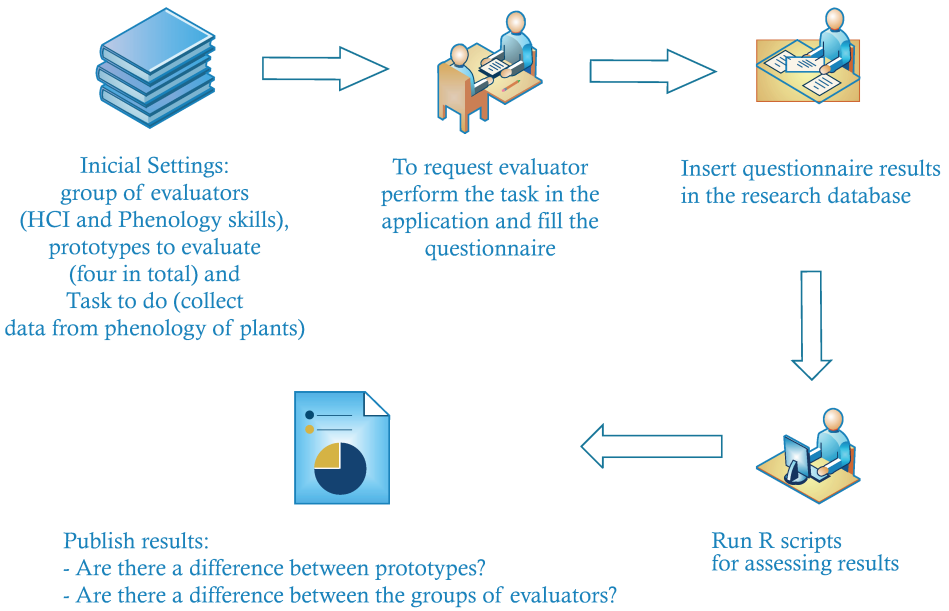


Figure 3.3: Evaluation process.

Chapter 4

Case Study

This chapter presents the conducted case study that aims to demonstrate the use of the proposed evaluation methodology and guidelines. This case study concerns the evaluation of the interfaces of four prototypes designed to support the phenological data acquisition process in the field. We first describe the data acquisition process scenario in Section 4.1. Next, we describe the evaluated prototypes in Section 4.2. Finally, in Section 4.3, we present and discuss the results obtained from the evaluation using the proposed methodology.

4.1 Phenology data acquisition in the field

Recently, phenology has been recognized as an important discipline for understanding the impact of climate change on living beings [21]. Phenology studies depend on the analysis of long-term temporal data. The common approach is the direct observation of plant individuals in the field at regular intervals (e.g., monthly or weekly) and the identification of phenophases (e.g., leafing, budding, flowering, and ripening) [22]. One widely adopted approach for data acquisition relies on using a qualitative method to assess the presence or absence of phenophase or using a quantitative method that assigns a different number (usually 0, 1, or 2) for a phenophase, depending on its intensity [25]. Usually, phenophase intensities are registered on paper sheets (in the field) and later inserted into digital spreadsheets (in the laboratory). This acquisition procedure can lead to errors and discrepancies in the collected data, which can delay data processing and analysis, as well as knowledge discovery.

Figure 4.1 presents the typical phenological data acquisition workflow. First, on-the-ground observations are planned. Multiple phenology experts may be involved in this process. Next, the in-the-field observations are performed by assigning intensity scores to plant phenophases. These scores are then registered in paper worksheets. In the lab, these data are stored in digital spreadsheets. At this moment, inserted data are checked with the objective of determining any inconsistency with previous annotations. If any inconsistency is identified, spreadsheets need to be updated accordingly.

In this context, we have been specifying and developing new applications to support data acquisition in the field, based on the ongoing phenological observations carried out

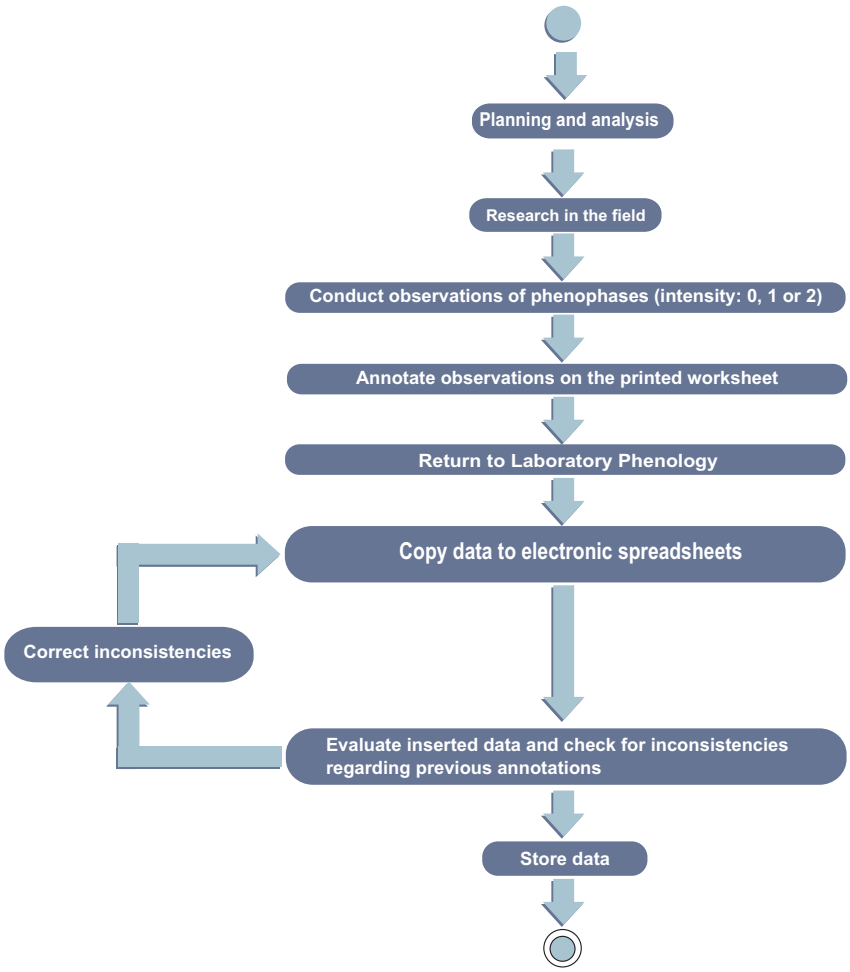


Figure 4.1: Typical data acquisition workflow. Figure adapted from Mariano [20].

by the group from Phenology Lab at UNESP.[1] The objective is to design and implement applications for portable devices that may support phenology experts in the field by: i) providing location-aware information regarding plant individuals; ii) monitoring the evolution of the data acquisition process on real time; and iii) implementing user friendly and loss-free mechanisms for data insertion and validation. The main challenges faced here rely on both the design and the in-the-field validation of appropriate interfaces for data insertion using portable devices, as well as the implementation of protocols to guarantee that no data are lost in the whole data acquisition process. This chapter addresses the interface design evaluation of developed prototypes using the proposed set of guidelines.

4.2 Evaluated Prototypes

The prototypes of phenological data acquisition applications considered in this study were object of design within the scope of a graduate course in HCI (second semester of 2012) at the Institute of Computing, University of Campinas, Brazil. The methodology used in the design process was proposed based on recent studies of usability and inspired by Participatory Design practices and the Organizational Semiotics theory [5].

[1]Details from field site and sample methods can be found elsewhere [2, 10].

The design problem proposed to the students involved support to activities the biologists develop both in the lab (Planning and Analysis) and in the field (Field Work). The Planning and Analysis are activities in the lab to prepare the field work (pre-field), monitoring its execution, and analysing data after field work (post-field). Thus, the designing problem involved (a) the design of a (web) application to support the planning of field work, and receiving and analyzing the data collected, and (b) the design of an application to support biologists in the field work. Both applications should communicate. The object of discussion in this dissertation is the mobile application to support the biologist field work.

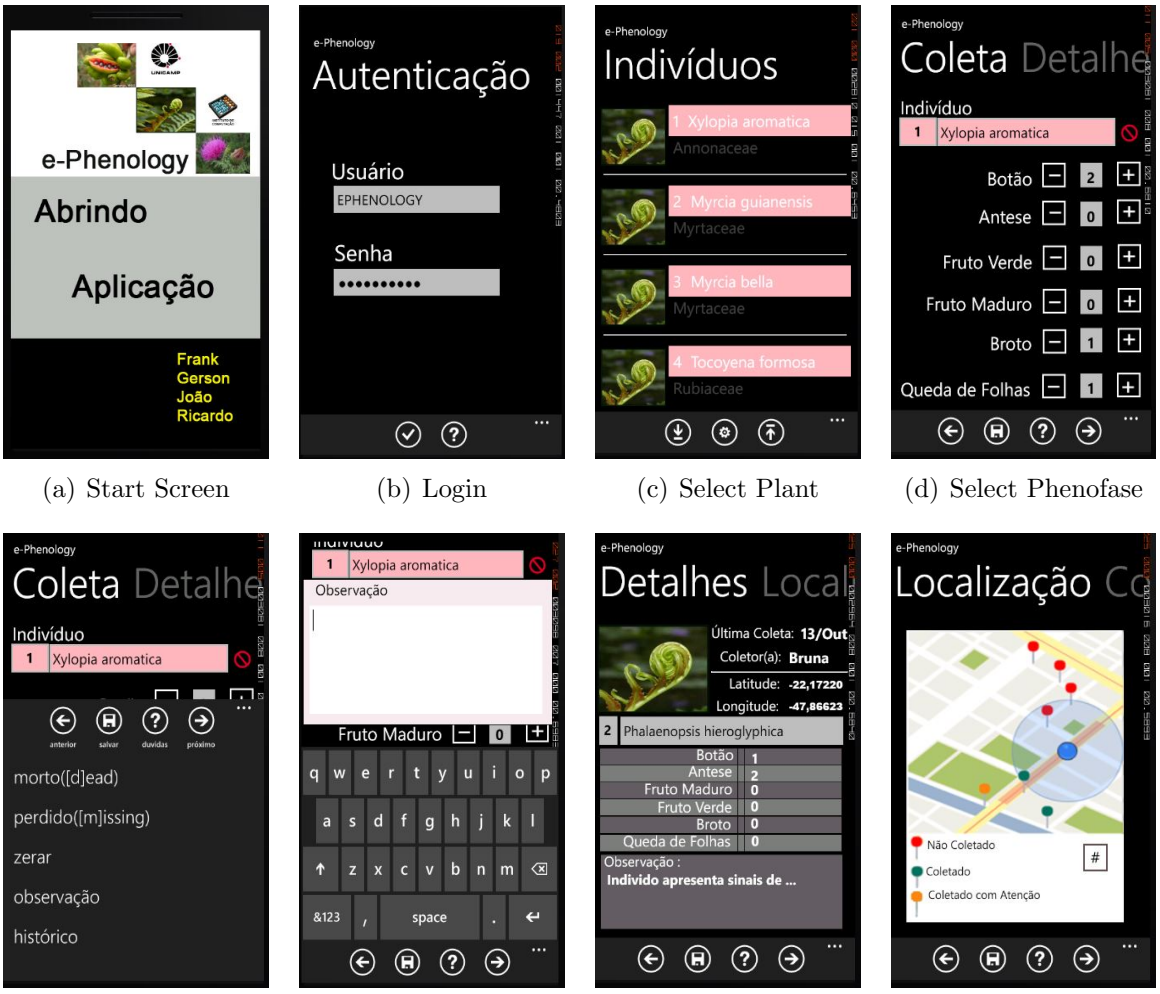
A set of 25 students, organized in 7 groups, worked in the role of designers to conceive and develop the interface of the application. Four groups designed mobile interfaces and three groups designed web applications for supporting the process management. All the groups conducted the following activities: i) problem clarification through participatory practices (e.g., Group Elicitation Method [9]) and context analysis through Organizational Semiotics' artifacts (e.g., Stakeholder Identification Diagram, Evaluation Framework [7]); ii) organization of a first set of requirements, prototyping (low and high fidelity) and evaluation in an iterative cycle.

During the process, the participants communicated with the partner biologists both on-line and in face-to-face meetings: from the very start when the problem was being clarified to the validation of requirements and the evaluation of different proposals. At the end of the term, the groups presented their prototypes to two biologists from the Phenology Laboratory and two Computer Scientist from the Institute of Computing, University of Campinas. They were very excited with the great possibilities of the prospective applications to facilitate and add to their work in data acquisition in the field. Our challenge was then to evaluate these prototypes with a sound set of guidelines in order to discover which one (or what aspects of them) would best fit to the needs of experts within this domain (ePhenology Project).

In Figures 4.2 to 4.5, we can see the screen shots of the prototypes considered in this evaluation. These screen shots refer to the main data acquisition process and some extra features implemented in each prototype. In the first prototype, Figure 4.2(d), phenophase scores are defined using the "minus" (–) and the "plus" (+) buttons. In the second prototype, Figure 4.3(c), the scores are defined using a sliding bar for each phenophase. In the third prototype, Figure 4.4(c), the scores are defined by clicking in the icon for each phenophase. Finally, in the fourth prototype, Figure 4.5(a), a quite different design is adopted, where phenophase scores are represented by painted icons: When only one icon is filled (see for example the leaf fall phenophase – *Queda* in Portuguese), then the intensity assigned to this phenophase is 1. When two icons are filled (see for example the flower bud phenophase – *Botão* in Portuguese), the intensity assigned is 2.

4.3 Results and Analysis

To evaluate the prototypes with the proposed guideline set, we invited a group of six specialists in HCI from the Institute of Computing, University of Campinas and six spe-



(a) Start Screen (b) Login (c) Select Plant (d) Select Phenofase

(e) Select Phenofase (f) Observations (g) Plant Details (h) Location Plant

Figure 4.2: Screen shots of Prototype 1.

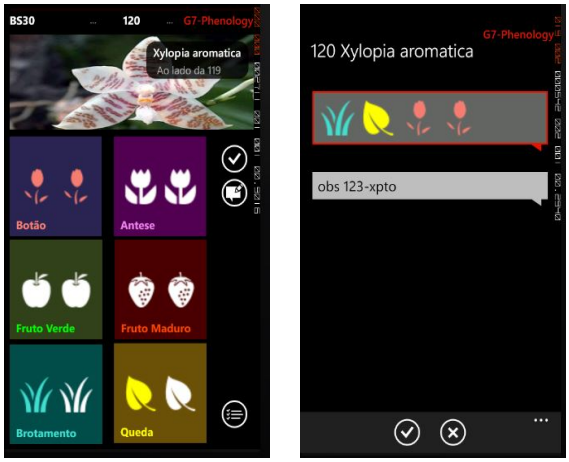
(a) Login (b) Select Local (c) Select Phenofase (d) Select Phenofase

Figure 4.3: Screen shots of Prototype 2.

cialists in Phenology from São Paulo State University – UNESP. For this evaluation, the proposed questionnaire was constructed from the guideline sentences and answered in a

(a) Login (b) Select Plant (c) Select Phenofase (d) Select Phenofase

Figure 4.4: Screen shots of Prototype 3.



(a) Select Phenofase (b) Select Phenofase

Figure 4.5: Screen shots of Prototype 4.

Likert scale (1-5). Based on the specialists' responses, we computed the average scores for each question. Figures 4.6, 4.7, 4.8, 4.9, 4.10, and 4.11 present these scores. Based on the evaluations, we highlight some important points:

- Regarding the Physical World step (see Figure 4.6), we can notice the high performance of all prototypes regarding Q6 ("Is it sufficiently easy to operate keys with one hand?"). We can also observe low average scores of all evaluated prototypes regarding Q3 ("Does the application provide voice assistance?"). Furthermore, it is worth mentioning the low performance of Prototype 2, with regard to Q4 ("The pictures on the screen are satisfactory quality and size") and the slight superior performance of Prototype 1 regarding Q5 ("Does the application provide appropriate menu button for touch screen?").

- Regarding the Empiric step (Figure 4.7), we can notice that all prototypes achieved relatively high average scores for Q7 ("Are the response time and information display

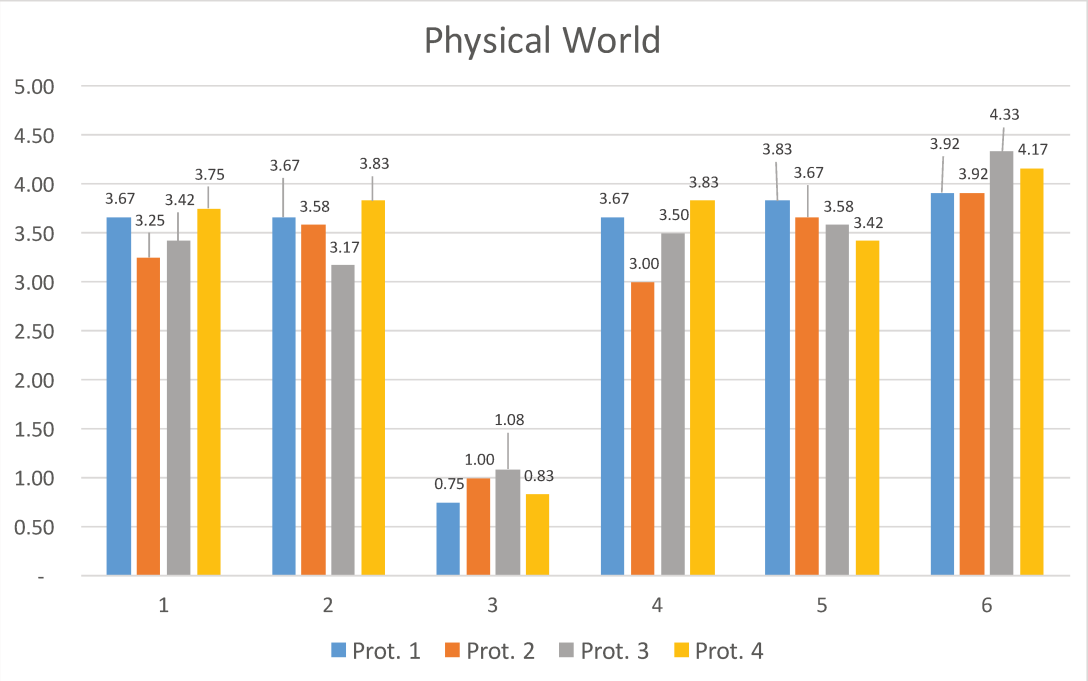


Figure 4.6: Average results for the Physical World step.

fast enough?") and low average scores for Q9 ("Are exchange and transmission of data between this product and other products (e.g., computer, PDA, and other mobile products) easy?"). The results for Q7 are due to the fact that we are working with prototypes. In this case, the response is very fast because there exist little hardware processing. For Q9, the low scores are due to the fact that this specific features was not implemented in the prototypes. Also, it is worth mentioning the low scores for Q10 ("Does the application provide automatic update?") observed for all prototypes, specially for Prototype 3.

- Regarding the Syntactic step (Figure 4.8), we can observe that Prototype 1 has the highest scores in 7 out of 12 questions. Another fact it is worth mentioning is that Prototype 4 has the lowest scores in 7 out of 12 questions. Also, we can notice that in questions Q11 ("How much information about system resources was displayed?") and Q19 ("Does application provide shortcuts for experienced users and wizards for new users?") all prototypes had scores below the media ("3.0 - Neither agree nor disagree") showing that this guideline was neglected by all of them.

- Regarding the Semantic step (Figure 4.9), we can observe that Prototype 2 is much better than the other ones for question Q25 ("Does the application provide appropriate help?"). It is also worth mentioning the low scores observed for Q23 ("Is the output data easy to use?") and for Q24 ("Are the documentation and manual for this application sufficiently informative?") for all prototypes. In Q26 ("Is feedback on the completion of tasks clear?"), Q27 ("Is the interface with this application clear and underatandable?") and Q28 ("Is the design of the graphic symbols, icons and labels on the icons sufficiently relevant?") all prototypes had scores above media except for Prototype 4 in question Q26.

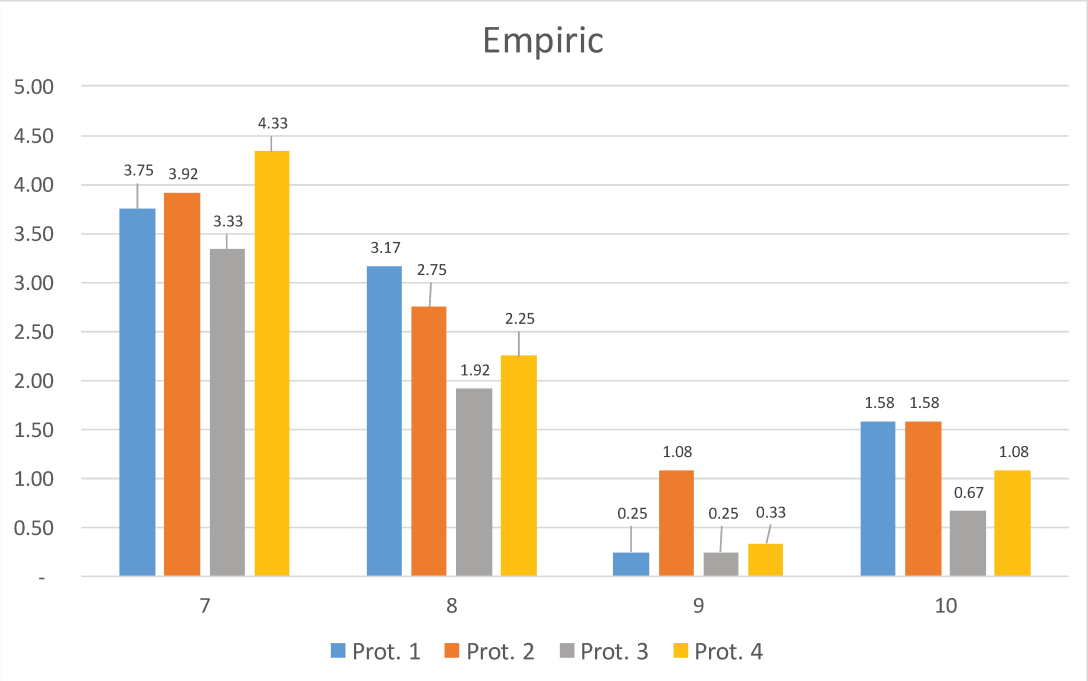


Figure 4.7: Average results for the Empiric step.

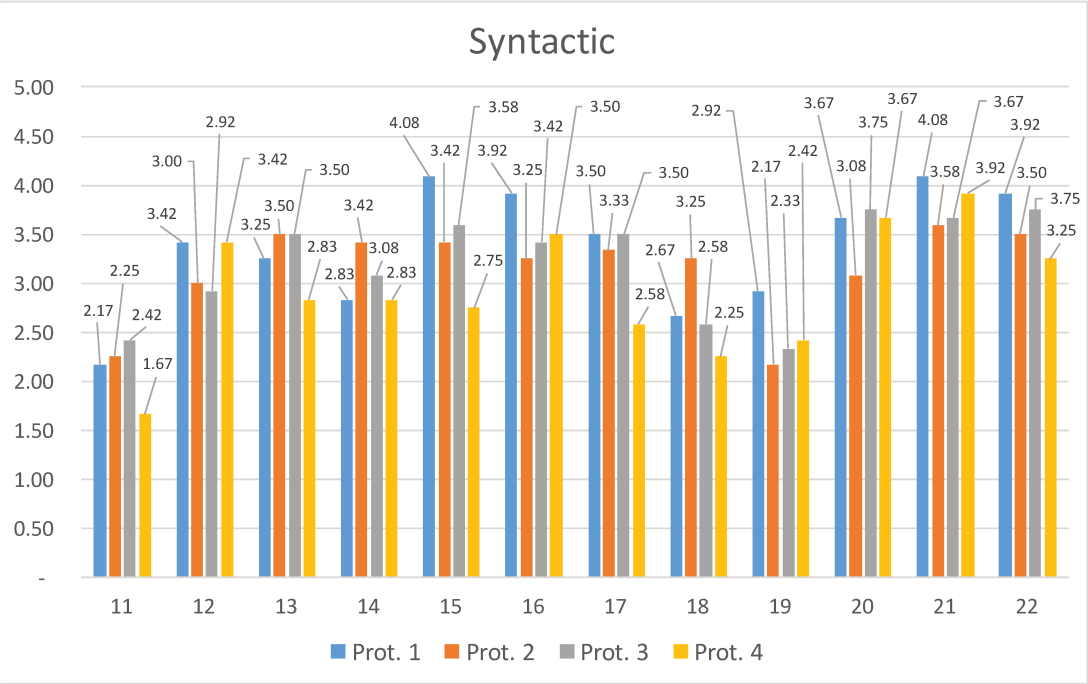


Figure 4.8: Average results for the Syntactic step.

- Regarding the Pragmatic step (Figure 4.10), there is no clear winner, except for Prototype 1 in Q33 ("Are the error messages effective in assisting you to fix problems?") and in Q36 ("Is it easy to learn to operate this application?"). It is worth mentioning the low scores observed for Q32 ("Does design of application is suitable for multiple contexts (home, business, travel, etc.) including support for runtime adaptation?") and Q35 ("Does application provide the ability to personalize the

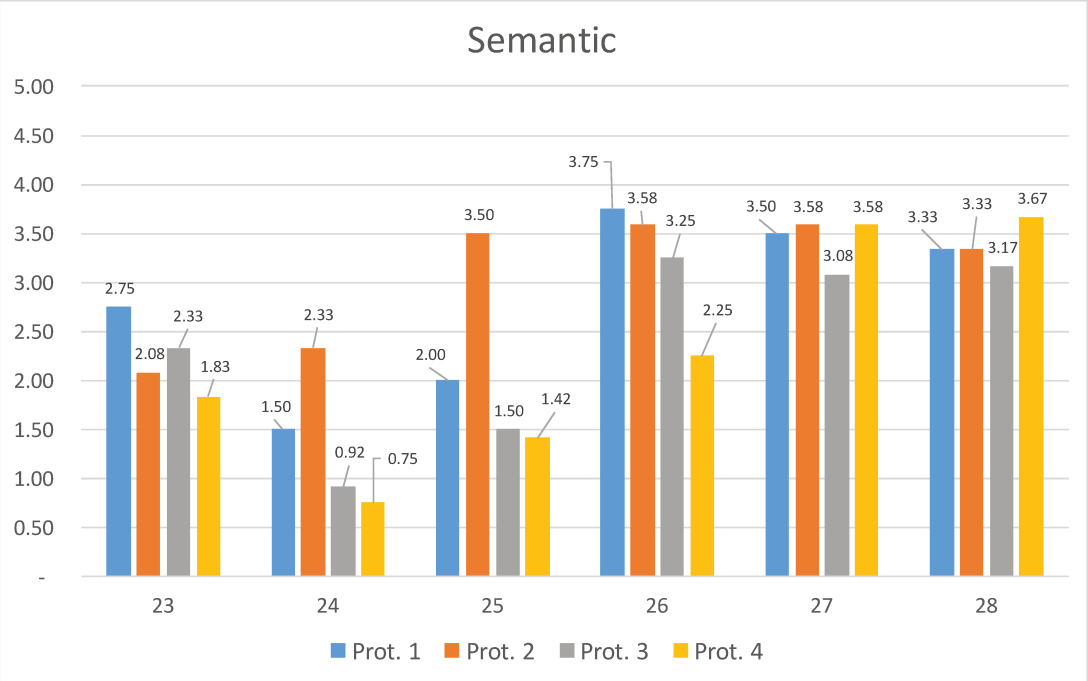


Figure 4.9: Average results for the Semantic step.

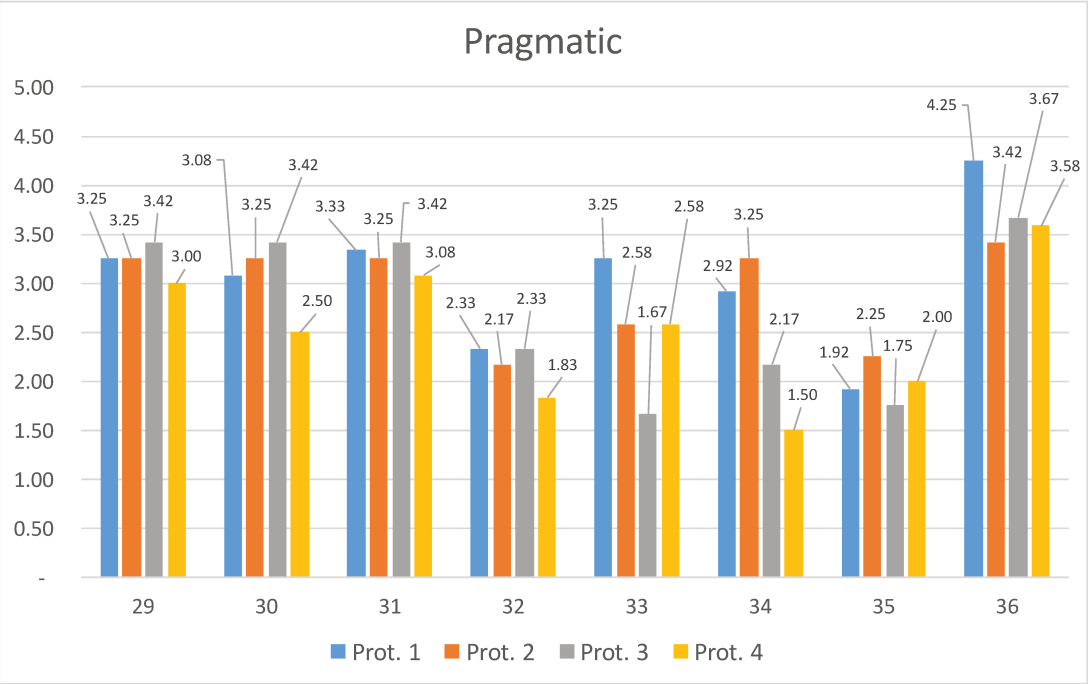


Figure 4.10: Average results for the Pragmatic step.

application to suit the user?") for all prototypes.

- Regarding the Social World step (Figure 4.11), all prototypes need to be improved. We can observe, however, that Prototype 4 has the worst score for Q37 ("Does application allow privacy and security control for single or multiple users?") and Q38 ("Does application have all the functions and capabilities you expect it to have?"), in which it had scores below average.

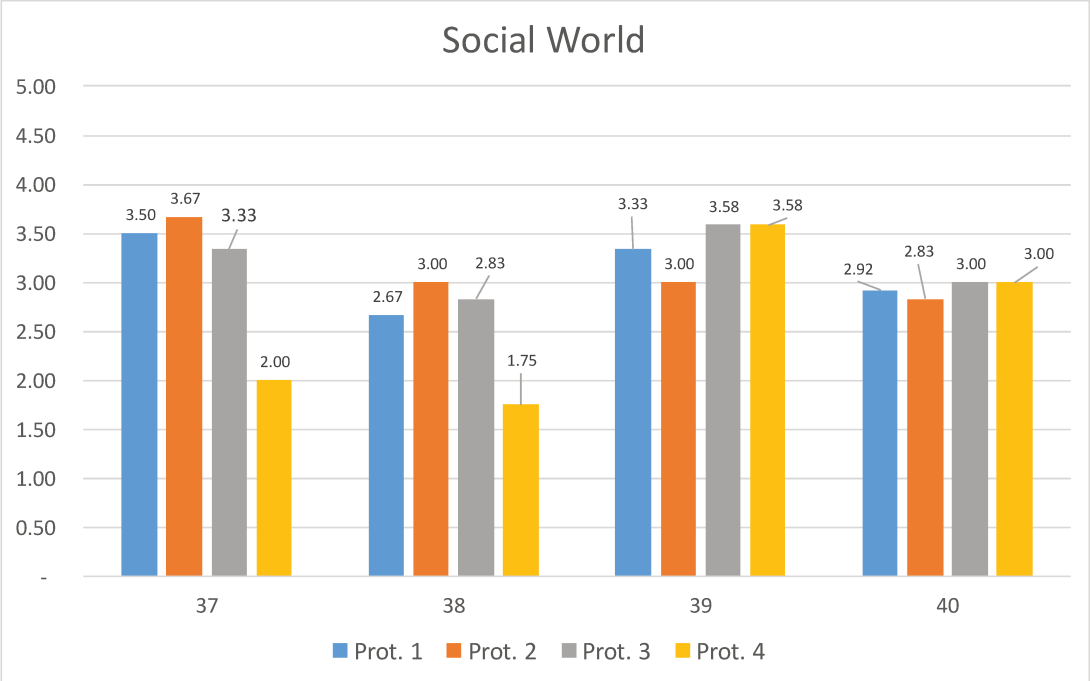


Figure 4.11: Average results for the Social World step.

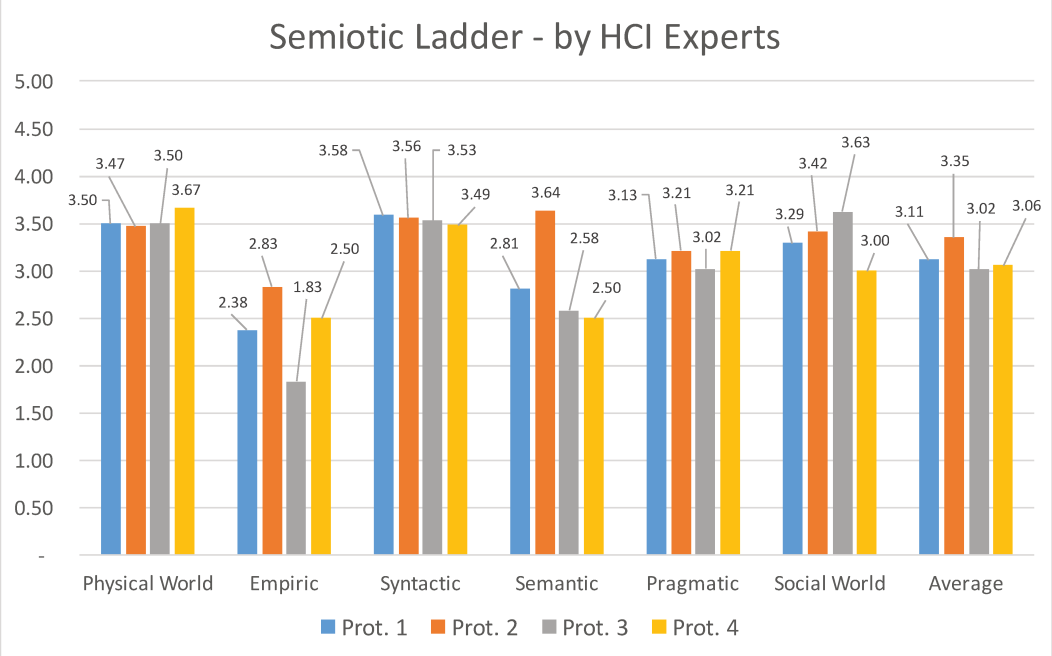


Figure 4.12: Average results for each Semiotic Ladder step according to the evaluation of HCI experts.

Figures 4.12, 4.13, and 4.14 show the average scores for each group of specialist and for each Semiotic Ladder step, considering its questions. As it can be observed, there is no clear winner prototype regarding all criteria. That is true by taking into account the evaluation of HCI experts and all evaluators altogether. For Phenology experts, we can observe that Prototype 1 has the highest scores.

Observing the results for HCI specialists (Figure 4.12), we can point out that Proto-

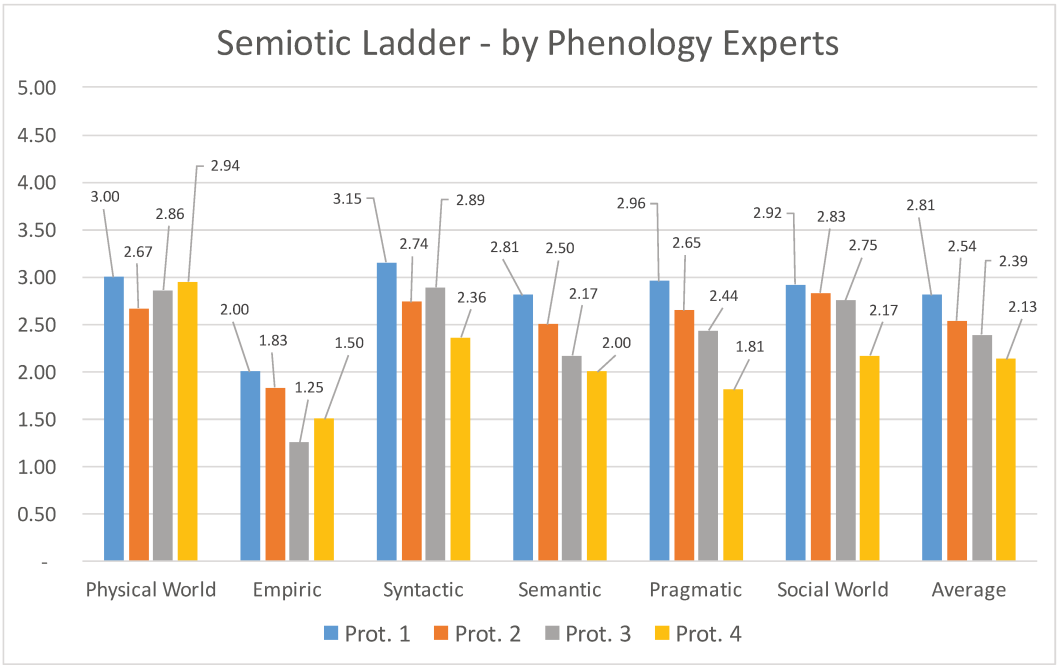


Figure 4.13: Average results for each Semiotic Ladder step according to the evaluation of Phenology experts.

type 1 is the best one in terms of the Syntactic step, while Prototype 2 is the best one in terms of the Empiric and Semantic steps. For the Physical World step, Prototype 4 is the best and for the Social World step, Prototype 3 has the highest average scores. For the Pragmatic step, we can notice that Prototypes 1, 2, and 4 have almost the same average scores.

Observing the results for Phenology specialists (Figure 4.13), we can point out that Prototype 1 is the best one for all steps. Prototype 2 is the second choice, having the highest scores (disregarding Prototype 1) for all steps except for the Physical World and Syntactic step, on which Prototype 3 has a better score (disregarding Prototype 1).

When we analyze the global result (Figure 4.14), i.e., considering both groups of evaluators altogether, we can observe that Prototype 1 is the best one in terms of the Syntactic and Pragmatic steps, while Prototype 2 is the best one in terms of the Empiric and Semantic steps. For the Physical World step, Prototype 4 is the best, while for the Social World step, Prototype 3 has the highest score.

If we consider that the minimum acceptable score should be 4 ("Agree") and that none of them reached this score, we conclude that all prototype must be improved to meet the standards of evaluators.

Figure 4.15 shows the evaluation scores of each prototype for each evaluator. Evaluators from 1 to 6 are HCI experts and 7 to 12 are Phenology experts. We can observe that Evaluator 3 (an HCI specialist) gave the best scores for all prototypes (except for Prototype 3, but it was almost the highest). We can also observe lowest average scores given by Evaluator 11 (a Phenology specialist). Another fact worth mentioning is that scores provided by HCI specialists are better than those given by Phenology specialists. This is probably because usually the HCI specialist is looking for a non-functional prototype of

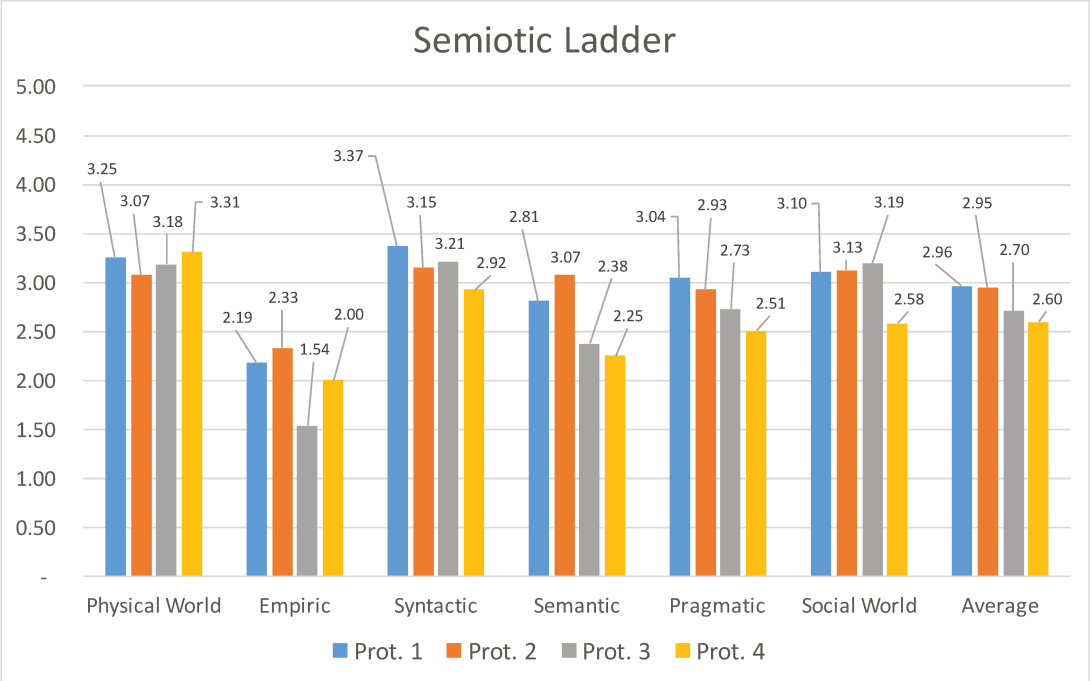


Figure 4.14: Average results for each Semiotic Ladder step considering both groups of evaluators (HCI and Phenology experts).

a prospective application and, on the other hand, the Phenology specialist is looking at a product that s/he wishes to use.

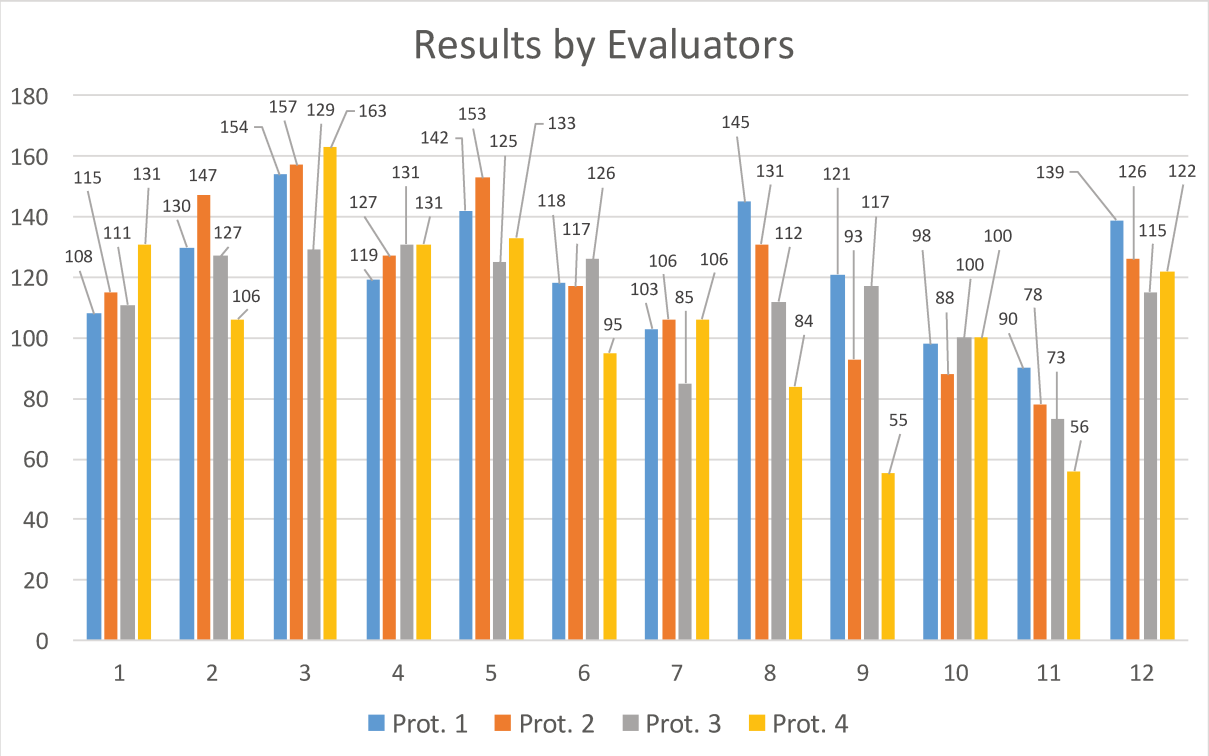


Figure 4.15: Total points obtained by each prototype for each evaluator.

Next, we present tables with another view of the results: by summing up the total scores of each prototype for each likert scale. In tables 4.1 to 4.7, we show the total points obtained for each prototype in a Likert scale, e.g, how many times each Lickert scale value was selected. Considering that 4 (agree) and 5 (strongly agree) states that the evaluator agrees that this guideline was covered by the prototype and 1 (completely disagree) and 2 (disagree) states that the Prototype do not, we provide a discussion regarding the adherence of each prototype considering a set of guidelines proposed for each step of the Semiotic Ladder.

Table 4.1: Total points by Likert Scale for each prototype in the Physical World step.

Physical World	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	5	8	4	17	23	15	12	38
Prototype 2	4	8	12	18	15	15	20	30
Prototype 3	4	7	10	13	27	11	17	38
Prototype 4	3	9	8	14	19	19	17	38

Table 4.1 shows the result for each prototype concerning with the Physical World step. Here, we note that Prototype 2 is the one with lowest overall score.

Table 4.2: Total points by Likert Scale for each prototype in the Empiric step.

Empiric	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	19	2	4	4	12	7	6	19
Prototype 2	18	1	2	10	8	9	3	17
Prototype 3	19	10	6	4	5	4	16	9
Prototype 4	21	2	3	8	6	8	5	14

Table 4.2 shows the result for each prototype in the Empiric step. Here, we note that many questions in this step were considered not applicable (value 0), probably because we are evaluating prototypes not completely functional. Also, we can note that Prototype 3 was the one that had the worst performance for this step.

Table 4.3 shows the result for each prototype in the Syntactic step. Here, we note that the Prototype 1 was the best and Prototype 4, the worst. However, all prototypes had a good evaluation for this step.

Table 4.4 shows the result for each prototype in the Semantic step. Here, we note that Prototype 2 was the best and Prototypes 3 and 4, the worst.

Table 4.5 shows the result for each prototype in the Pragmatic step. In this step, we can see that Prototype 1 had the best and at the same time the worst scores. This occurs

Table 4.3: Total points by Likert Scale for each prototype in the Syntactic step.

Syntactic	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	11	4	16	27	62	24	20	86
Prototype 2	13	8	14	39	50	20	22	70
Prototype 3	8	8	19	38	53	18	27	71
Prototype 4	19	17	16	27	34	31	33	65

Table 4.4: Total points by Likert Scale for each prototype in the Semantic step.

Semantic	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	13	2	9	16	26	6	11	32
Prototype 2	12	0	7	18	22	13	7	35
Prototype 3	11	11	13	17	17	3	24	20
Prototype 4	15	14	11	10	14	8	25	22

Table 4.5: Total points by Likert Scale for each prototype in the Pragmatic step.

Pragmatic	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	8	6	20	13	38	11	26	49
Prototype 2	8	8	10	32	33	5	18	38
Prototype 3	10	10	14	32	22	8	24	30
Prototype 4	21	9	12	19	24	11	21	35

because Prototypes 2 and 3 had a lot of neutral scores (3), while Prototype 4 had many questions with no applicable evaluation (0).

Table 4.6 shows the result for each prototype in the Social World step. Here, we can note that Prototype 1, 2 and 3 had good scores, while Prototype 4 had more negative scores (1 and 2) than positive (4 and 5).

In Table 4.7, we show the total points obtained for each prototype. Considering this global view, we point out that:

- Prototype 1 is the most adhered to the proposed guideline followed by 2 then 3 and finally 4.
- Prototype 2 is the one that had more neutral evaluations.

Table 4.6: Total points by Likert Scale for each prototype in the Social World step.

Social World	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	4	1	9	13	14	7	10	21
Prototype 2	3	1	8	16	15	5	9	20
Prototype 3	3	0	10	11	20	4	10	24
Prototype 4	6	8	9	9	9	7	17	16

Table 4.7: Total points by Likert Scale for each prototype.

Prototype	0	1	2	3	4	5	1 + 2	4 + 5
Prototype 1	60	23	62	90	175	70	85	245
Prototype 2	58	26	53	133	143	67	79	210
Prototype 3	55	46	72	115	144	48	118	192
Prototype 4	85	59	59	87	106	84	118	190

- Prototype 4 is the one that had more not applicable evaluation but, also, it has more maximums scores (5). The not applicable evaluation could be explained because is the prototype with less features implemented.

Based on all these analyses, we raise the following issues to be observed in the design of a new interface phenology data acquisition mobile application:

- Consider the aspects that have been neglected by Prototype 2 regarding the questions of the physical world step.
- Consider the aspects that led Prototypes 1, 3 and 4 to have good ratings in physical world step.
- Analyze the questions that have zero value in the empiric step.
- Consider the aspects that had good evaluations in the Syntactic step.
- Observe what Prototypes 3 and 4 neglected in the Semantic step and why Prototypes 1 and 2 had good results in this same step.
- For the Pragmatic step, analyze why Prototype 1 had the best and the worst scores and, also, why Prototypes 2 and 3 had many neutral scores (3).
- Analyze what aspects Prototype 4 do not cover in the social world step and look for good features of the the other prototypes concerning this step.

Furthermore, we also recommend to take into account the comments made by the phenology experts evaluators, which are reproduced in Appendix D.

The results of the evaluation through the proposed guidelines show that all the prototypes have interesting design decisions to be considered in the design of a final product. These results are even more important because they came from different design proposals that were created based on an informed and well-defined design process conducted by prospective designers in a participatory style.

4.4 Statistical Analysis

Based on the results of the conducted evaluation, statistical analyses were performed in order to establish if there are differences among the evaluations of the different prototypes and also if there are differences between the two groups of evaluators (specialists in HCI and specialists in Phenology).

These analyses were done using the evaluation methodology proposed in Section 3.2.1. Following this methodology, we characterize the performed statistical tests.

4.4.1 Overview

1. Definition of groups of evaluators.
   - Group 1: Six HCI experts from the Institute of Computing, University of Campinas;
   - Group 2: Six Phenology Experts from the São Paulo State University – UNESP.
2. Definition of application interfaces.
   - Four prototypes developed for in-the-field phenological data acquisition (for more details, see Section 4.2).
3. Definition of hypotheses:

   The performed analysis addresses two hypotheses:

   (a) Are there differences between the groups of evaluators considering
      - their general evaluation of the prototypes?
      - their evaluation of a particular prototype?
      - their evaluation of prototypes using a particular Semiotic Ladder step?
      - their evaluation of a particular prototype according to a specific Semiotic Ladder step?

   (b) Are there differences among the prototypes considering
      - the complete set of guideline questions, i.e., all Semiotic Ladder steps altogether?
      - a particular Semiotic Ladder step?
4. Definition of statistical tests: based on characteristics of these questions, statistical tests was selected according to Table 3.8
   - For the first set of questions, we conduct t tests;
   - For the second set of questions, we use ANOVA Repeated Measures.

4.4.2 Statistical Results

For each prototype, we sum up the scores assigned by each evaluator. The final scores are then used in our statistical tests. In Tables A.1 to A.4, we present the final scores obtained by each prototype in the evaluations performed with specialists in HCI and in Phenology.

In Table 4.8, we present the results for the first set of questions, while in Table 4.9 for the second set of questions. In these tables, "Yes" means that the null hypothesis was rejected, i.e., there are differences among the groups of evaluators (Table 4.8), or among the evaluated prototypes (Table 4.9). In all tests, we consider 95% confidence. We can observe that both groups of evaluators have different views regarding all prototypes (first column in Table 4.8). Furthermore, they also have different opinions regarding all prototypes (except for Prototype 1), when all steps are considered (first line). In fact, both groups of evaluators have the same opinion for all steps, regarding Prototype 1. For Prototype 2, there is statistical difference only for the Semantic step and when considering all steps. For Prototypes 3 and 4, sometimes they agree (for example, for the evaluation of the Physical World and Semantic steps), while for other steps they disagree (Syntactic and Pragmatic Step). With regard to the Empiric step, they have the same opinion with regard to Prototype 3, but, for Prototype 4 they have different points of view.

Table 4.8: Result of t tests – Are there differences among the groups of evaluators?

•	All Prototypes	Prototype 1	Prototype 2	Prototype 3	Prototype 4
All Steps	Yes	No	Yes	Yes	Yes
Physical World	Yes	No	No	No	No
Empiric	Yes	No	No	No	Yes
Syntactic	Yes	No	No	Yes	Yes
Semantic	Yes	No	Yes	No	No
Pragmatic	Yes	No	No	Yes	Yes
Social World	Yes	No	No	No	No

Regarding the prototypes, there are statistical differences when the scores provided of HCI and Phenology experts are considered altogether (first column in Table 4.9). However, when only the opinion of HCI experts are considered, no differences are identified, regardless the Semiotic Ladder step (second column). For Phenology experts, the prototypes are different when all steps are considered altogether, or when only the Semantic step is taken into account.

Finally, in Table 4.10, we present the statistical results concerning the identification of the best prototypes, considering the opinion of different groups of experts and different Semiotic Ladder steps. We can observe that when all steps are considered as well as the evaluation of all evaluators, Prototype 1 is the best one. For HCI experts, there is not a clear winner regardless the Semiotic Ladder step (cells marked with "$*$"). For Phenology experts, Prototype 1 is the best one when all steps are considered altogether, or when only the Semantic step is taken into account.

Table 4.9: Results of ANOVA tests – Are there differences among the prototypes?

•	In General	HCI experts	Phenology experts
All Steps	Yes	No	Yes
Physical World	Yes	No	No
Empiric	Yes	No	No
Syntactic	Yes	No	No
Semantic	Yes	No	Yes
Pragmatic	Yes	No	No
Social World	Yes	No	No

Table 4.10: Which one is the best Prototype?

•	In General	HCI experts	Phenology experts
All Steps	1	*	1
Physical World	4	*	*
Empiric	2	*	*
Syntactic	1	*	*
Semantic	2	*	1
Pragmatic	1	*	*
Social World	3	*	*

Chapter 5

Conclusion

The large-scale adoption of portable device applications depends on the use of careful interface design. In this dissertation, we analyzed strategies, guidelines, and questions in literature on mobile application design and evaluation and proposed a novel set of guidelines composed of 27 semiotic-informed guidelines for support the evaluation of mobile application interfaces. This guideline set served also as the basis for the proposal of a methodology for evaluating mobile application interfaces. Furthermore, we implemented a set of tools to support the adoption of the proposed methodology/guidelines in future evaluations. Another contribution of this work refers to the fact that the proposed method can be applied to people working in the application domain, without specialized knowledge on semiotics or on methods for interface evaluation.

We demonstrate the use of the proposed method and guidelines in the context of the evaluation of four prototypes recently proposed for phenological data acquisition. The analysis of results from evaluations indicates that the proposed guidelines set is well suited for the evaluation of mobile application interface as it helps to identify positive and negative aspects of proposed designs, according to well-defined semiotic concepts of different information layers.

Part of the results described in this work has been reported in a paper accepted for publication in the 17th International Conference on Enterprise information Systems (ICEIS 2015),[1] which took place in Barcelona in April, 2015.

A limitation of this study is the fact of having conducted the experiments with a not totally functional application. In this case, some questions are not applicable or not completely applicable, and this fact could have affected the results of our study. Example of questions not applicable include: Does the application provide voice assistance? Does the application provide automatic update? Another point that is worth mentioning is the fact that not all guidelines are applicable to any mobile devices. Some of them are specific for small devices, such as smart phones. An example is Guideline 4 - Easy operation with one hand. Also, some questions could not be applicable to some applications, e.g. Question 2 - "Does the application provide a virtual keypad?".

[1] Nicastro, F. ; Pereira, R. ; Alberton, B. ; Morellato, L. P. C. ; Baranauskas, M. C. ; Torres, R. da S. . A Semiotic-Informed Approach to Interface Guidelines for Mobile Applications: A Case Study on Phenology Data Acquisition. In: 17th International Conference on Enterprise information Systems, 2015, Barcelona, Spain. ICEIS, 2015. To appear.

As future work, we intend to develop novel interface designs based on the findings related to the advantages and drawbacks identified in the evaluated prototypes, and to conduct experiments in which phenology experts will be able to evaluate the developed prototypes using the proposed guidelines. We also intend to evaluate whether the guidelines set proposed in this dissertation are also useful for supporting the design of mobile application interfaces. We do believe that the set of guidelines is generic enough to be used for different application domains. Additionally, we are planning to conduct further evaluation activities to assess the guideline's contributions in different design contexts and with different groups of experts.

Bibliography

[1] M. Abdallah. Home healthcare devices: Towards a scalable, portable, accurate, and affordable data acquisition instrument. In *38th Annual Northeast Bioengineering Conference (NEBEC)*, pages 9–10, March 2012.

[2] B. Alberton, J. Almeida, R. Helm, R. da S. Torres, A. Menzel, and L. P. C. Morellato. Using phenological cameras to track the green up in a cerrado savanna and its on-the-ground validation. *Ecological Informatics*, 19(0):62 – 70, 2014.

[3] J. Almeida, J. A. Dos Santos, B. Alberton, R. da S. Torres, and L. P. C. Morellato. Remote phenology: Applying machine learning to detect phenological patterns in a cerrado savanna. In *8th International Conference on E-Science*, pages 1–8, Oct 2012.

[4] M. C. C. Baranauskas. O modelo semioparticipativo de design. In *Codesign de Redes Digitais - Tecnologia e Educação a Serviço da Inclusão Social*, pages 38–66. Penso Editora Ltda, 2013.

[5] M. C. C. Baranauskas. Social awareness in HCI. *Interactions*, 21(4):66–69, July 2014.

[6] M. C. C. Baranauskas and H. V. da Rocha. Design e avaliação de interfaces humano-computador. *Campinas–SP: Nied/Unicamp*, 2003.

[7] M. C. C. Baranauskas, J. Schimiguel, C. A. C. Simoni, and C. M. B. Medeiros. Guiding the process of requirements elicitation with a semiotic approach. In *11th International Conference on Human-Computer Interaction*, pages 100–111, 2005.

[8] N. Z. binti Ayob, Ab. R. C. Hussin, and H. M. Dahlan. Three layers design guideline for mobile application. In *International Conference on Information Management and Engineering (ICIME)*, pages 427–431, April 2009.

[9] G. A. Boy. The group elicitation method for participatory design and usability testing. *interactions*, 4(2):27–33, 1997.

[10] M. G. G. Camargo, R. M. Souza, P. Reys, and L. P. C Morellato. Effects of environmental conditions associated to the cardinal orientation on the reproductive phenology of the cerrado savanna tree xylopia aromatica (annonaceae). *Biotropica*, 83(3):1007–1019, 2011.

[11] J. Cordeiro and J. Filipe. The semiotic pentagram framework–a perspective on the use of semiotics within organisational semiotics. In *Proceedings of the 7th International Workshop on Organisational Semiotics*, pages 249 – 265, 2004.

[12] H. B.-L. Duh, G. C. B. Tan, and V. H. Chen. Usability evaluation for mobile device: A comparison of laboratory and field tests. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '06, pages 181–186, 2006.

[13] J. Gong and P. Tarasewich. Guidelines for handheld mobile device interface design. In *Proceedings of DSI 2004 Annual Meeting*, pages 3751–3756, 2004.

[14] J. Hao, S. H. Kim, S. A. Ay, and R. Zimmermann. Energy-efficient mobile video management using smartphones. In *Proceedings of the second annual ACM conference on Multimedia systems*, pages 11–22, 2011.

[15] A. Hussain and M. Kutar. Usability metric framework for mobile phone application. *PGNet*, pages 978–1, 2009.

[16] C. D. Keeling and T. P. Whorf. Increased activity of northern vegetation inferred from atmospheric CO2 measurements. *Nature*, 382(6587):146–149, 1996.

[17] J. Lazar, J. H. Feng, and H. Hochheiser. *Research methods in human-computer interaction*. John Wiley & Sons, 2010.

[18] K. Liu. *Semiotics in information systems engineering*. Cambridge University Press, 2000.

[19] D. Loustau, A. Bosc, A. Colin, J. Ogée, H. Davi, C. François, E. Dufrêne, M. Déqué, E. Cloppet, D. Arrouays, et al. Modeling climate change effects on the potential production of french plains forests at the sub-regional level. *Tree Physiology*, 25(7):813–823, 2005.

[20] G. C. Mariano. Especificação e implementação do banco de dados do projeto e-phenology. Master's thesis, Instituto de Computação, Universidade Estadual de Campinas, 2013.

[21] A. Menzel, T. H. Sparks, N. Estrella, E. Koch, A. Aasa, R. Ahas, K. Alm-kübler, P. Bissolli, O. Braslavská, A. Briede, et al. European phenological response to climate change matches the warming pattern. *Global change biology*, 12(10):1969–1976, 2006.

[22] L. P. C. Morellato, M. G. G. Camargo, F. F. D. Neves, B. G. Luize, A. Mantovani, and I. L. Hudson. The influence of sampling method, sample size, and frequency of observations on plant phenological patterns and interpretation in tropical forest trees. In Irene L. Hudson and Marie R. Keatley, editors, *Phenological Research*, pages 99–121. 2010.

[23] F. Nayebi, J. M. Desharnais, and A. Abran. The state of the art of mobile application usability evaluation. In *25th Canadian Conference on Electrical Computer Engineering (CCECE)*, pages 1–4, April 2012.

[24] G. C. S. Negi. Leaf and bud demography and shoot growth in evergreen and deciduous trees of central himalaya, india. *Trees*, 20(4):416–429, 2006.

[25] F. F. D. Neves and L. P. C. Morellato. Métodos de amostragem e avaliação utilizados em estudos fenológicos de florestas tropicais. *Acta Botanica Brasileira*, 18(1):99–108, 2004.

[26] C. Parmesan and G. Yohe. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918):37–42, 2003.

[27] L. S. G. Piccolo and M. C. C. Baranauskas. Climbing the ladder with energy: Informing the design of eco-feedback technology with a social approach. *The Fourteenth International Conference on Informatics and Semiotics in Organisations*, pages 187–194, 2013.

[28] N. Radia, Y. Zhang, M. Tatipamula, and V. K. Madisetti. Next-generation applications on cellular networks: Trends, challenges, and solutions. 100(4):841–854, April 2012.

[29] B. Rathcke and E. P. Lacey. Phenological patterns of terrestrial plants. *Annual Review of Ecology and Systematics*, pages 179–214, 1985.

[30] M. Rauch. Mobile documentation: Usability guidelines, and considerations for providing documentation on kindle, tablets, and smartphones. In *Professional Communication Conference (IPCC)*, pages 1–13, Oct 2011.

[31] P. B. Reich. Phenology of tropical forests: patterns, causes, and consequences. *Canadian Journal of Botany*, 73(2):164–174, 1995.

[32] A. D. Richardson, B. H. Braswell, D. Y. Hollinger, J. P. Jenkins, and S. V. Ollinger. Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications*, 19(6):1417–1428, 2009.

[33] C. Rosenzweig, D. K., M. Vicarelli, P. Neofotis, Q. Wu, G. Casassa, A. Menzel, T. L. Root, N. Estrella, B. S., et al. Attributing physical and biological impacts to anthropogenic climate change. *Nature*, 453(7193):353–357, 2008.

[34] T. Rötzer, R. Grote, and H. Pretzsch. The timing of bud burst and its effect on tree growth. *International Journal of Biometeorology*, 48(3):109–118, 2004.

[35] Y. S. Ryu. *Development of usability questionnaires for electronic mobile products and decision making methods*. PhD thesis, Virginia Polytechnic Institute and State University, 2005.

[36] M. D. Schwartz. *Phenology: An Integrative Environmental Science*. Springer Netherlands, 2013.

[37] S. B. Shneiderman and C. Plaisant. Designing the user interface 4 th edition. *ed: Pearson Addison Wesley, USA*, 2005.

[38] R. Stamper. *Information in Business and Administrative Systems*. John Wiley & Sons, Inc., New York, NY, USA, 1973.

[39] G. R. Walther. Plants in a warmer world. *Perspectives in plant ecology, evolution and systematics*, 6(3):169–185, 2003.

[40] G. R. Walther, E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. Ecological responses to recent climate change. *Nature*, 416(6879):389–395, 2002.

[41] I. Zamzami and M. Mahmud. User satisfaction on smart phone interface design, information quality evaluation. In *International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, pages 78–82, Nov 2012.

[42] I. P. Zarko, A. Antonic, and K. Pripužic. Publish/subscribe middleware for energy-efficient mobile crowdsensing. In *Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 1099–1110, 2013.

Appendix A

Evaluations

This appendix shows the form filled out by evaluators (Figure A.1) and the overall scores obtained for each prototype and Semiotic Ladder step (Tables A.1 to A.4).

Name of Reviewer: _____ Age:_____

Literacy: _____

Evaluated Application: Interface Application's Prototype for e-Phenology Project. Group: ___

Put a "x" in the column that best describe your feeling in relation the evaluated application. Use this scale:
1 - Strongly disagree; 2 Disagree; 3 -Neither agree nor disagree; 4 Agree; 5 - Strongly agree.
Leave blank if you think the question does not apply in this context

Q.No.	Questions	1	2	3	4	5
	Physical World					
1	The backlighting feature for the keyboard and screen is appropriate in all contexts.					
2	The application provides a functional virtual keypad.					
3	The application provides voice assistance.					
4	The pictures on the screen are satisfactory quality and size.					
5	The application provides appropriate menu button for touch screen.					
6	It is sufficiently easy to operate keys with one hand.					
	Empiric					
7	The response time and information display are fast enough.					
8	Recent data items are kept during a short time.					
9	Exchange and transmission of data between this product and other products (e.g., computer, PDA, and other mobile products) are easy.					
10	Application provides automatic update.					
	Syntactic					
11	Enough information about system resources are displayed.					
12	HOME and MENU buttons are sufficiently easy to locate for all operations.					
13	Letter codes for the menu selection are designed carefully.					
14	Color coding and data display are compatible with familiar conventions.					
15	All operations can be carried out in a systematically similar way.					
16	The organization of information on the application screen is clear.					
17	It is easy to navigate between hierarchical menus, pages, and screen.					
18	This application enables the quick, effective, and economical performance of tasks.					
19	The application provides feedback (haptic, audio, visual, etc.) constantly in order to keep the user engaged and attentive.					
20	The application provides shortcuts for experienced users and wizards for new users.					
21	Interact with this product don't require a lot of mental effort.					
22	It is easy for you to remember how to perform tasks with this product.					
	Semantic					
23	Feedback on the completion of tasks is clear.					
24	The output data are easy to use.					
25	The documentation and manual for this application are sufficiently informative.					
26	The application provides appropriate help.					
27	The interface with this application is clear and understandable.					
28	The design of the graphic symbols, icons and labels on the icons are sufficiently relevant.					
	Pragmatic					
29	It is easy to access the information that you need from the application.					
30	Discovering new features is sufficiently easy.					
31	The application allows convenient use with the ability to handle multiple and frequent interruptions with limited attention from the user.					
32	The design of application is suitable for multiple contexts (home, business, travel, etc.) including support for runtime adaptation.					
33	The error messages are effective in assisting you to fix problems.					
34	The messages aimed at prevent you from making mistakes are adequate.					
35	The application provides the ability to personalize its enviroment to suit the user.					
36	It is easy to learn to work with this application.					
	Social World					
37	The application allows privacy and security control for single and multiple users.					
38	The application has all functions and capabilities you expected it to had.					
39	This application is attractive and pleasing.					
40	The application is secure to use on moving as when driving or walking.					

What is your oppinion about this guideline?

_____

_____

If you want, write comments, criticism and suggestions about the prototypes:

_____

_____

_____

_____

Figure A.1: Questionnaire based on the proposed guideline.

Table A.1: Overall scores for Prototype 1.

Ev.	Expertise	Total	PW	Emp	Syn	Sem	Prag	SW
1	HCI	108	21	6	33	16	22	10
2	HCI	130	25	10	42	18	22	13
3	HCI	154	22	17	54	19	28	14
4	HCI	119	19	8	39	15	24	14
5	HCI	142	23	7	46	20	30	16
6	HCI	118	16	9	44	13	24	12
7	Phenology	103	21	3	36	7	24	12
8	Phenology	145	24	15	41	18	32	15
9	Phenology	121	18	4	41	24	20	14
10	Phenology	98	14	6	32	16	20	10
11	Phenology	90	14	9	33	15	15	4
12	Phenology	139	17	11	44	21	31	15

Table A.2: Overall scores for Prototype 2.

Ev.	Expertise	Total	PW	Emp	Syn	Sem	Prag	SW
1	HCI	115	18	5	40	19	24	9
2	HCI	147	26	15	45	23	21	17
3	HCI	157	20	16	49	28	30	14
4	HCI	127	22	6	37	21	27	14
5	HCI	153	24	16	50	21	27	15
6	HCI	117	15	10	35	19	25	13
7	Phenology	106	16	4	38	12	21	15
8	Phenology	131	21	10	42	14	30	14
9	Phenology	93	19	4	18	20	21	11
10	Phenology	88	10	6	29	12	21	10
11	Phenology	78	12	10	28	13	12	3
12	Phenology	126	18	10	42	19	22	15

Table A.3: Overall scores for Prototype 3.

Ev.	Expertise	Total	PW	Emp	Syn	Sem	Prag	SW
1	HCI	111	21	8	39	15	19	9
2	HCI	127	20	7	40	19	24	17
3	HCI	129	20	7	44	18	24	16
4	HCI	131	23	5	44	14	28	17
5	HCI	125	21	7	46	12	26	13
6	HCI	126	21	10	41	15	24	15
7	Phenology	85	14	1	29	8	23	10
8	Phenology	112	20	4	38	16	21	13
9	Phenology	117	24	4	45	11	20	13
10	Phenology	100	14	6	34	16	18	12
11	Phenology	73	14	4	25	14	12	4
12	Phenology	115	17	11	37	13	23	14

Table A.4: Overall scores for Prototype 4.

Ev.	Expertise	Total	PW	Emp	Syn	Sem	Prag	SW
1	HCI	131	22	7	46	19	26	11
2	HCI	106	25	10	30	12	17	12
3	HCI	163	24	16	54	24	32	13
4	HCI	131	26	8	48	11	28	10
5	HCI	133	21	11	41	11	34	15
6	HCI	95	14	8	32	13	17	11
7	Phenology	106	18	5	39	10	24	10
8	Phenology	84	15	6	22	15	17	9
9	Phenology	55	21	4	14	10	3	3
10	Phenology	100	20	6	27	15	18	14
11	Phenology	56	14	4	24	8	4	2
12	Phenology	122	18	11	44	14	21	14

Appendix B

Tools and Scripts

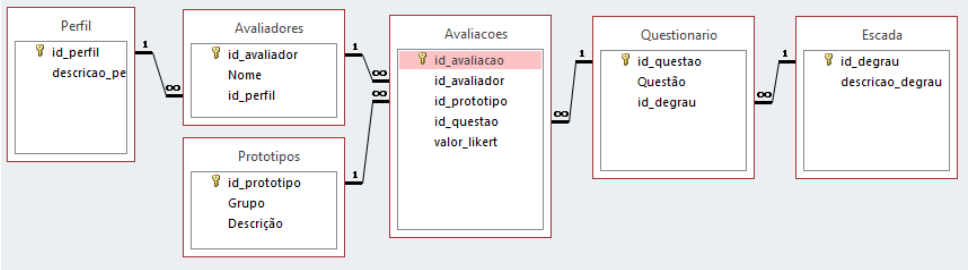


Figure B.1: Entity-relationship model of the database used to store evaluation scores.

- Table Avaliacoes

```
CREATE TABLE [dbo].[Avaliacoes](
       [id_avaliacao] [int] NOT NULL,
       [id_avaliador] [int] NULL,
       [id_prototipo] [int] NULL,
       [id_questao] [int] NULL,
       [valor_likert] [int] NULL
) ON [PRIMARY]
GO
```

- Table Avaliadores

```
CREATE TABLE [dbo].[Avaliadores](
       [id_avaliador] [int] NOT NULL,
       [Nome] [nvarchar](255) NULL,
       [id_perfil] [int] NULL
) ON [PRIMARY]
GO
```

- Table Escada

```
CREATE TABLE [dbo].[Escada](
       [id_degrau] [int] NOT NULL,
       [descricao_degrau] [nvarchar](255)
NULL
) ON [PRIMARY]
GO
```

- Table Perfil

```
CREATE TABLE [dbo].[Perfil](
       [id_perfil] [int] NOT NULL,
       [descricao_perfil] [nvarchar](255)
NULL
) ON [PRIMARY]
GO
```

- Table Prototipos

```
CREATE TABLE [dbo].[Prototipos](
       [id_prototipo] [int] NOT NULL,
       [Grupo] [int] NULL,
       [Descrição] [nvarchar](255) NULL
) ON [PRIMARY]
GO
```

- Table Questionario

```
CREATE TABLE [dbo].[Questionario](
       [id_questao] [int] NOT NULL,
       [Questão] [nvarchar](max) NULL,
       [id_degrau] [int] NULL
) ON [PRIMARY] TEXTIMAGE_ON [PRIMARY]
GO
```

Figure B.2: Script to create a database to store evaluation scores.

```
SELECT [Avaliacoes Consulta].Nome, [Avaliacoes Consulta].id_prototipo,
        [Avaliacoes Consulta].Descrição, [Avaliacoes Consulta].id_degrau,
        [Avaliacoes Consulta].descricao_degrau, [Avaliacoes Consulta].id_perfil,
        [Avaliacoes Consulta].descricao_perfil, Sum([Avaliacoes Consulta].valor_likert) AS Total
FROM [Avaliacoes Consulta]
GROUP BY    [Avaliacoes Consulta].Nome, [Avaliacoes Consulta].id_prototipo,
            [Avaliacoes Consulta].Descrição, [Avaliacoes Consulta].id_degrau,
            [Avaliacoes Consulta].descricao_degrau, [Avaliacoes Consulta].id_perfil,
            [Avaliacoes Consulta].descricao_perfil
ORDER BY    [Avaliacoes Consulta].id_degrau, [Avaliacoes Consulta].descricao_degrau,
            [Avaliacoes Consulta].id_perfil, [Avaliacoes Consulta].descricao_perfil,
            [Avaliacoes Consulta].Nome, [Avaliacoes Consulta].Descrição;
```

Figure B.3: Example of a query for retrieving data of interest.

```
library(RODBC)
channel <- odbcConnect("bdAvaliacoes")
# Load the data by evaluator profile and by steps of semiotic ladder
# Step 1 - Physycal World
GeralStep1IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 1")
GeralStep1Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 1")
Prot1Step1IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 1 and id_prototipo =1")
Prot1Step1Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 1 and id_prototipo =1")
Prot2Step1IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 1 and id_prototipo =2")
Prot2Step1Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 1 and id_prototipo =2")
Prot3Step1IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 1 and id_prototipo =3")
Prot3Step1Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 1 and id_prototipo =3")
Prot4Step1IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 1 and id_prototipo =4")
Prot4Step1Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 1 and id_prototipo =4")
# Step 2 - Empiric
GeralStep2IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 2")
GeralStep2Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 2")
Prot1Step2IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 2 and id_prototipo =1")
Prot1Step2Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 2 and id_prototipo =1")
Prot2Step2IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 2 and id_prototipo =2")
Prot2Step2Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 2 and id_prototipo =2")
Prot3Step2IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 2 and id_prototipo =3")
Prot3Step2Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 2 and id_prototipo =3")
Prot4Step2IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 2 and id_prototipo =4")
Prot4Step2Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 2 and id_prototipo =4")
# Step 3 - Syntactic
GeralStep3IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 3")
GeralStep3Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 3")
Prot1Step3IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 3 and id_prototipo =1")
Prot1Step3Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 3 and id_prototipo =1")
Prot2Step3IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 3 and id_prototipo =2")
Prot2Step3Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 3 and id_prototipo =2")
Prot3Step3IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 3 and id_prototipo =3")
Prot3Step3Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 3 and id_prototipo =3")
Prot4Step3IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 3 and id_prototipo =4")
Prot4Step3Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 3 and id_prototipo =4")
# Step 4 - Semantic
GeralStep4IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 4")
GeralStep4Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 4")
Prot1Step4IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 4 and id_prototipo =1")
Prot1Step4Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 4 and id_prototipo =1")
Prot2Step4IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 4 and id_prototipo =2")
Prot2Step4Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 4 and id_prototipo =2")
Prot3Step4IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 4 and id_prototipo =3")
Prot3Step4Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 4 and id_prototipo =3")
Prot4Step4IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 4 and id_prototipo =4")
Prot4Step4Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 4 and id_prototipo =4")
# Step 5 - Pragmatic
GeralStep5IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 5")
GeralStep5Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 5")
Prot1Step5IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 5 and id_prototipo =1")
Prot1Step5Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 5 and id_prototipo =1")
Prot2Step5IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 5 and id_prototipo =2")
Prot2Step5Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 5 and id_prototipo =2")
Prot3Step5IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 5 and id_prototipo =3")
Prot3Step5Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 5 and id_prototipo =3")
Prot4Step5IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 5 and id_prototipo =4")
Prot4Step5Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 5 and id_prototipo =4")
# Step 6 - Social World
GeralStep6IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 6")
GeralStep6Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 6")
Prot1Step6IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 6 and id_prototipo =1")
Prot1Step6Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 6 and id_prototipo =1")
Prot2Step6IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 6 and id_prototipo =2")
Prot2Step6Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 6 and id_prototipo =2")
Prot3Step6IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 6 and id_prototipo =3")
Prot3Step6Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 6 and id_prototipo =3")
Prot4Step6IHC <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'IHC' and id_degrau = 6 and id_prototipo =4")
Prot4Step6Unesp <- sqlQuery(channel, "select Total from conPorDegrauDetalhado where descricao_perfil = 'Unesp' and id_degrau = 6 and id_prototipo =4")
```

Figure B.4: Example of an R script for retrieving data of interest.

```
# Execute t test for each pair HCI/Unesp
# Physical World
t.test(GeralStep1IHC, GeralStep1Unesp)
t.test(Prot1Step1IHC, Prot1Step1Unesp)
t.test(Prot2Step1IHC, Prot2Step1Unesp)
t.test(Prot3Step1IHC, Prot3Step1Unesp)
t.test(Prot4Step1IHC, Prot4Step1Unesp)
# Empiric
t.test(GeralStep2IHC, GeralStep2Unesp)
t.test(Prot1Step2IHC, Prot1Step2Unesp)
t.test(Prot2Step2IHC, Prot2Step2Unesp)
t.test(Prot3Step2IHC, Prot3Step2Unesp)
t.test(Prot4Step2IHC, Prot4Step2Unesp)
# Syntactic
t.test(GeralStep3IHC, GeralStep3Unesp)
t.test(Prot1Step3IHC, Prot1Step3Unesp)
t.test(Prot2Step3IHC, Prot2Step3Unesp)
t.test(Prot3Step3IHC, Prot3Step3Unesp)
t.test(Prot4Step3IHC, Prot4Step3Unesp)
# Semantic
t.test(GeralStep4IHC, GeralStep4Unesp)
t.test(Prot1Step4IHC, Prot1Step4Unesp)
t.test(Prot2Step4IHC, Prot2Step4Unesp)
t.test(Prot3Step4IHC, Prot3Step4Unesp)
t.test(Prot4Step4IHC, Prot4Step4Unesp)
# Pragmatic
t.test(GeralStep5IHC, GeralStep5Unesp)
t.test(Prot1Step5IHC, Prot1Step5Unesp)
t.test(Prot2Step5IHC, Prot2Step5Unesp)
t.test(Prot3Step5IHC, Prot3Step5Unesp)
t.test(Prot4Step5IHC, Prot4Step5Unesp)
# Social World
t.test(GeralStep6IHC, GeralStep6Unesp)
t.test(Prot1Step6IHC, Prot1Step6Unesp)
t.test(Prot2Step6IHC, Prot2Step6Unesp)
t.test(Prot3Step6IHC, Prot3Step6Unesp)
t.test(Prot4Step6IHC, Prot4Step6Unesp)
```

Figure B.5: Example of an R script for executing t tests.

Appendix C

Spreadsheets Used for In-the-field Data Acquisition

			Coletor:		Data:							
Loca	Parce	Indivíd	Família	Espécie	Bot	Ante	FV	FM	Brot	Quec	Indiv	Observações
BL	2	89	Verbenaceae	*Aegiphila lhotszkyana*							89	bem na borda
BL	2	90	Fabaceae	*Anadenanthera falcata*							90	
BL	2	91	Fabaceae	*Bauhinia rufa*							91	
BL	2	92	Myrtaceae	*Campomanesia pubescens*							92	
BL	2	93	Myrtaceae	*Campomanesia pubescens*							93	
BL	2	94	Styracaceae	*Styrax ferrugineus*							94	
BL	2	95	Rubiaceae	*Palicourea rigida*							95	perdida mai/07
BL	2	96	Fabaceae	*Dalbergia miscolobium*							96	
BL	2	97	Ochnaceae	*Ouratea spectabilis*							97	
BL	2	98	Nyctaginaceae	*Guapira opositta*							98	
BL	2	100	Malpighiaceae	*Byrsonima intermedia*							100	
BL	2	101	Fabaceae	*Dalbergia miscolobium*							101	
BL	2	102	Fabaceae	*Stryphnodendron adstringens*							102	depois da 106
BL	2	104	Fabaceae	*Stryphnodendron obovatum*							104	bem baixa, frente da 102, morta?
BL	2	105	Fabaceae	*Bowdichia virgilioides*							105	Perdida?
BL	2	106	Lauraceae	*Ocotea pulchella*							106	
BL	2	107	Lauraceae	*Ocotea pulchella*							107	
BL	2	109	Fabaceae	*Anadenanthera falcata*							109	placa alta, lado da 118
BL	2	110	Malpighiaceae	*Byrsonima intermedia*							110	bem baixa, pé da 111, 60cm
BL	2	111	Erythroxylaceae	*Erythroxylum suberosum*							111	morta jan/2010
BL	2	112	Connaraceae	*Connarus suberosus*							112	
BL	2	113	Ochnaceae	*Ouratea spectabilis*							113	
BL	2	114	Erythroxylaceae	*Erythroxylum tortuosum*							114	morta jan/2010
BL	2	115	Bignoniaceae	*Tabebuia ochracea*							115	pé 120, morta?
BL	2	117	Asteraceae	*Gochnatia barrosii*							117	lado da 118, 109 e 110
BL	2	118	Malpighiaceae	*Byrsonima coccolobifolia*							118	s/ placa
BL	2	119	Malpighiaceae	*Byrsonima basiloba*								

Figure C.1: Example of a spreadsheet where in-the-field plant observations are registered.

			Coletor:		Data:							
Local	Trans	Individu	Família	Espécie	Bot	Ante	FV	FM	Brot	Quec	Indiv	Observações
BL	01 Bc	1	Vochysiaceae	Qualea grandiflora							1	s/ placa
BL	1	2	Myrtaceae	Myrcia guianensis							2	
BL	1	3	Myrtaceae	Myrcia bella							3	
BL	1	4	Rubiaceae	Tocoyena formosa							4	
BL	1	5	Rubiaceae	Tocoyena formosa							5	placa no chão
BL	1	7	Annonaceae	Xylopia aromatica							7	sem placa, atrás da 5
BL	1	8	Verbenaceae	Aegiphila lhotszkyana							8	sem placa, bem baixa, à direita da 5
BL	1	9	Annonaceae	Xylopia aromatica							9	
BL	1	10	Annonaceae	Xylopia aromatica							10	
BL	1	11	Fabaceae	Anadenanthera falcata							11	ao lado da placa 10
BL	1	12	Annonaceae	Xylopia aromatica							12	
BL	1	13	Malpighiaceae	Banisteriopsis campestris							13	
BL	1	14	Araliaceae	Schefflera vinosa							14	ver base
BL	1	15	Fabaceae	Machaerium acutifolium							15	lado da 14
BL	1	16	Ochnaceae	Ouratea spectabilis							16	baixa, pé 14
BL	1	18	Annonaceae	Xylopia aromatica							18	
BL	1	22	Fabaceae	Bauhinia rufa							22	
BL	1	23	Sapotaceae	Pouteria ramiflora							23	
BL	1	24	Sapotaceae	Pouteria ramiflora							24	
BL	1	25	Melastomataceae	Miconia rubiginosa							25	
BL	1	26	Bignoniaceae	Tabebuia ochracea							26	perdida abr/07

Figure C.2: Digital spreadsheet into which observation data are inserted after field work.

Appendix D

Comments by Phenology Experts Evaluators

- Evaluator 7
  - Prototype 1: *A forma de coleta de dados é boa, é fácil ir com (+) e (-) para marcar fenofase. Da informação de indivíduo já observado. Não encontrei como voltar a tela inicial da parcela com a lista de indivíduos depois que já comecei a coletar. Não há campo para informação de morto ou perdido, somente há o campo de observação. Há um bug no final quando chegamos no último indivíduo, começa a vibrar e cai o aplicativo.*
  - Prototype 2: *A aplicação possui praticamente todas as opções necessárias de observações como individuo perdido, morto etc. Bem completa e separa os ambientes, dá histórico do indivíduo etc. Achei que faltou um campo de observação para escrever, o botão de anotações parece não funcionar. O campo com o número do individuo pode ser editado, isso não é bom, podemos perder esta informação. Eu não gosto muito da forma de quantificar a fenofase. Mas o aplicativo traz bastante informação. Só achei que o aplicativo não é muito intuitivo, por exemplo, para encontrar todas as opções de informação do indivíduo o símbolo é (...), poderia ser um sinal de (+); deste jeito quase me passou despercebido. Acho que é o aplicativo mais completo em informação para o trabalho de campo.*
  - Prototype 3: *O Aplicativo possui as informações que desejamos, porém é um pouco lento para passarmos para o proximo individuo. As informações que editamos não foram salvas. O aplicativo permite ir adiante e voltar sempre, porém achei muito complexo de informações, e a forma de coletar também não me agradou. O principal é que é lento e não gostei das telas. Ao iniciar ele trava e não continua caso ele esteja no modo avião, acusa problemas com o GPS.*
  - Prototype 4: *É intuitiva com o uso dos desenhos para as fenofases. Acho mais fácil o usuário errar o valor da fenofase caso de um clique duplo sobre o desenho. A ferramenta de escrever observação para individuo não funciona.*

Não consegui rever os dados coletados. Acho que faltou suporte para os casos onde perdemos ou temos indivíduos mortos.

- Evaluator 8
  - Prototype 1: *Achei bem organizado e fácil de mexer. As informações dos indivíduos são de fácil acesso e aparece a lista de todos, o que facilita o trabalho. A forma de inserir os dados é semelhante àquela usada no trabalho normal de campo (planilhas) o que facilita bastante sua utilização. As informações são bem expostas e fáceis de acessar. A informação da observação anterior (dados coletados no mês anterior) são importantes de serem facilmente recuperadas. Uma mudança importante seria termos que adicionar apenas a fenofase que esta ocorrendo (as presenças) não sendo necessários adicionar os zeros quando a fenofase estiver ausente.*
  - Prototype 2: *Achei seguro e de fácil manipulação. Achei importante a escolha das áreas antes de iniciar as coletas ou escolher os indivíduos. Isso evitaria alguns erros. Mas a exposição dos dados anteriores (status se o individuo esta morto por exemplo) apesar de ser de fácil acesso não é muito evidente na tela principal (tela de atualização e inserção dos dados). A informação da fenologia anterior e da atual não ficaram muito claras para mim. Achei um pouco confuso.*
  - Prototype 3: *Achei fácil, porém esta com muitos erros. O erro do GPS foi constante e isto, com certeza, será um problema no campo. O sinal não deveria limitar a entrada no aplicativo. As informações não estão sendo salvas (aparece como coletado mas com a informação não salva). Aparecer a foto se coletado ou não é bem importante. O modo de salvar é demorado, poderia permitir salvar sem perguntar para cada indivíduo. Talvez uma opção de salvar por parcela seja mais rápida (exemplo salva-se as informações de cada individuo, aparece como coletado e após o término da parcela aparece a opção de salvar e pergunta se quer salvar as informações da parcela X).*
  - Prototype 4: *Achei pouco informativa e pouco direta. A maneira como se insere os dados da fenologia é pouco informativa e não muito clara. Acredito que a visualização dos valores (0, 1, 2) é um meio mais direto e informativo do que apenas uma visualização de cores. As cores em campo, podem ser difíceis de ver, o que pode levar a muitos erros na coleta dos dados. Uma maneira mais facil de encontrar qual individuo e área se deseja coletar seria importante (não consegui escolher um indivíduo aleatoriamente).*
- Evaluator 9
  - Prototype 1: *Ter que preencher sempre com zero as fenofases é ruim.*
  - Prototype 2: *Muito simples de usar, só a forma de classificar as fenofases que não achei prática.*
  - Prototype 3: *Problema com GPS e não salva a coleta.*

  - Prototype 4: *Esquema de marcação de fenofases muito bom, mas pode ser ruim com luz do sol no campo.*

- Evaluator 10

  - Prototype 1: *Deixar o zero já marcado e as espécies que já estão perdidas também. Dar um feedback se há espécies sem preencher.*
  - Prototype 2: Without Comments.
  - Prototype 3: *Não salva as alterações; bom que esta mostrando o que já foi desenvolvido; fator positivo ter a localização do individuo e as fotos.*
  - Prototype 4: Without Comments.

- Evaluator 11

  - Prototype 1: Needs to be improved for being usefull - very Slow! The default needs to be zero, otherwise it's very time consuming. We only note when there is a phenophase plus or equal 1. Queda can be 100 percent; you need to alow this possibility. Pictures can not be changed.
  - Prototype 2: Not very usefull for field work. The only think I like of this appplication is that you can see the plot and how much has already been sampled. There are mistakes with the names of species. You cannot see the full list of species in a whole.
  - Prototype 3: Needs a lot of improvments for being usefull: SLOW! It's not saving information. GPS is giving problemns: the app is blocked! The spacial separation of individuals in different plots is not clear: you need to separate by plots.
  - Prototype 4: Not usefull for field work.

- Evaluator 12

  - Prototype 1: *Comparando o G7 com o G1, o marcador das fenofases acho mais interessante ser composto por números que por figuras. As cores das figuras podem não ficar muito evidentes no sol. Quando o símbolo de exclamação aparece ao lado da espécie, o indivíduo não fica marcado como completo, era interessante que também ficasse verde para mostrar que já foi feita a avaliação.*
  - Prototype 2: *Comparando este aplicativo com o G7 e o G1, acho que este poderia ser composto por números (as fenofases) e não por barras. A visualização da coleta anterior achei um pouco confusa. Gostei da ideia de colocar a localização do ponto de coleta na tela inicial. Não gostei de todos os passos ter a opção de vibrar, e não conseguir retirar a vibração. Precisa ter os nomes corretos como "interior sul" e não "inferior sul".*
  - Prototype 3: *Achei bem simples, mas acho que poderia modificar a configuração dos números das fenofases, para facilitar visualizar no sol. Não tinha as informações da ultima coleta. Não esta fixando as informações e nem as fotos.*

– Prototype 4: *Como é o aplicativo mais simples, falta estar completo para poder avaliar melhor. No sol, acho que as figuras podem ficar um pouco difíceis de visualizar o que esta marcado.*