Investigating agreement among HCI researchers about Human-Computer Interaction in CMMI-DEV model

Taisa Guidini Gonçalves, Christophe Kolski, Pierre Loslever, Káthia Marçal de Oliveira

Author version, published in: *Interacting With Computers*, 32 (1), pp. 81-100, 2020.

Investigating agreement among HCI researchers about Human-Computer Interaction in CMMI-DEV model: a case study

Taisa Guidini Gonçalves, Pierre Loslever, Káthia Marçal de Oliveira and Christophe Kolski

LAMIH CNRS UMR 8201, Université Polytechnique Hauts-de-France, Le Mont Houy, 59313 Valenciennes Cedex 9, France
taisaguidini@gmail.com, pierre.loslever@uphf.fr, kathia.oliveira@uphf.fr, christophe.kolski@uphf.fr

Abstract. Do HCI researchers agree about which HCI approaches should be used in the development of an interactive system? To address this question, we performed a study based on the engineering process areas (requirements development, technical solution, product integration, verification, and validation) of the CMMI-DEV (Capability Maturity Model Integration for Development), a software engineering model commonly used in industry. To carry out this study, a literature review was performed, a specific instrument was designed based on CMMI-DEV, and interviews with twenty researchers from the HCI domain were conducted. Analyzing the interview data with monovariate and multivariate statistical approach (Multiple Correspondence Analysis - MCA), we find the greatest agreement occurred among researchers for HCI approaches that support verification and validation phases, and the use of functional prototypes in some activities of the last phases of software development (technical solution and product integration). However, we identified lack of agreement among researchers regarding requirements development activities.

Keywords: Human-Computer Interaction; Multiple Correspondence Analysis; Interactive system; CMMI-DEV; Case study.

Research Highlights

- Integration of HCI approaches in the CMMI-DEV model for interactive systems generated useful information for software developers.
- Interviews with 20 HCI researchers from the HCI domain provided useful qualitative data.
- Analysis of 37 questions using monovariate and multivariate statistical approaches provided interesting results.
- Results of this study reveal a range of agreement among researchers regarding HCI approaches.

1. Introduction

The development of interactive systems is supported by approaches (methods, techniques, patterns, and standards) that focus on Human-Computer Interaction (HCI) issues (for instance, design issues (Schutte, 2017) (Shin et al., 2017), solutions, and applications for HCI (Sears & Jacko, 2009); approaches to deal with usability issues (Seffah et al., 2005) (Ham, 2014) (Mohamed et al., 2017); HCI standards (for instance, ISO 9241-11:2018 (ISO, 2018), ISO 9241-112:2017 (ISO, 2017a), ISO 9241-125:2017 (ISO, 2017b), ISO 9241-161:2016 (ISO, 2016), ISO 9241-151:2008 (ISO, 2008), and ISO/TR 16982:2002 (ISO, 2002); and usability capability/maturity models (Jokela et al., 2006) (Ogunyemi, Lamas, & Eze, 2018)). Despite this, the literature shows that HCI approaches and practices are not or are insufficiently used in the industry (see for instance (Bevan, 2009) (Hao & Jaafar, 2011)

(Scheiber et al., 2012) (Ogunyemi, Lamas, Adagunodo, Loizides, & Rosa, 2016) (Ogunyemi et al., 2018)).

Taking advantage of the use of software process capability maturity (SPCM) models in the industry, we conducted a first research study aiming to identify which HCI approaches could support the software engineering practices of the Capability Maturity Model Integration for Development – CMMI-DEV (CMMI Product Team, 2010). To that end, we performed interviews with several HCI researchers. A detailed analysis of this study and the final set of HCI approaches to support CMMI-DEV can be found in (Gonçalves et al., 2016) (Gonçalves et al., 2018). This study presents a perspective on CMMI-DEV engineering processes from the point of view of the state of the art in HCI. It is important to emphasize that this work is focused on software aspects; it does not consider hardware aspects.

When analyzing the data from the interviews to identify the HCI approaches proposed by the researchers, we find a range of opinions. We decided to perform a second study and re-analyze the same data collected from the interviews with a different goal: to investigate the range of opinion amongst respondents. Unlike the first study, we were not interested in the suggested HCI approaches; but the level of agreement between the HCI researchers. For better confidence in the results, we decided to perform this second study supported by a well-founded statistical analysis. This paper presents the results of this second study.

We start this paper, by briefly introducing the main concepts of CMMI-DEV model (section 2) and the first study (Gonçalves et al., 2018) about the identification of HCI approaches to support CMMI-DEV (section 3). In section 4, we describe the detailed information about the instrument and the statistical methods. Section 5 describes the statistical analysis that we performed to investigate the agreement among HCI researchers, and a discussion about the results. Then, section 6 presents the threats to validity of this study. Finally, in section 7, we present our conclusions and research perspectives.

2. Software process capability maturity model: CMMI-DEV

CMMI-DEV is a workbench of software engineering best practices organized in different elements as follows:

- **Process area** - the core element composed of a set of practices related to an area divided in 1 to 3 specific goals.
- **Specific Goal (SG)** - objectives of improvement in each process area.
- **Specific Practice (SP)** - software engineering best practices to a single process area. A set of SP composes a SG.

CMMI-DEV version 1.3 (CMMI Product Team, 2010) has 22 process areas organized into four categories: engineering, support, project management, and process management.

In both studies, we focused on the **engineering category** since it covers the whole software development life cycle. The engineering category of CMMI-DEV is composed of five process areas: Requirements Development (RD); Technical Solution (TS); Product Integration (PI); Verification (VER); and Validation (VAL). Our study was performed considering the specific goals and practices (Figure 1) for these five process areas.

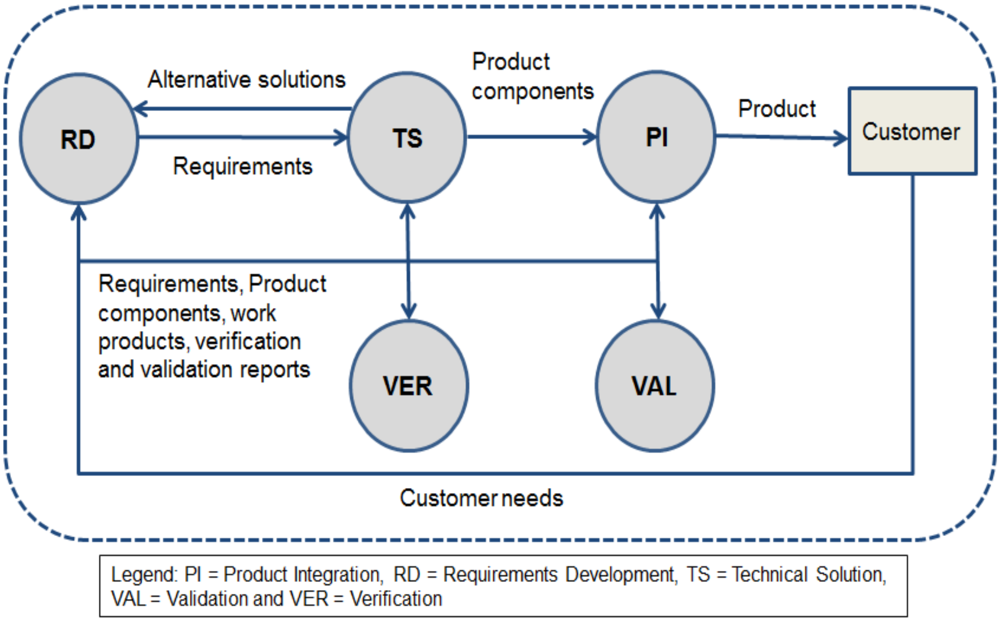

Figure 1. Engineering process areas (Gonçalves et al., 2016)

3. Supporting CMMI for developers with HCI approaches

To identify HCI approaches that could support the CMMI-DEV practices, we followed three main phases presented in Figure 2.

In the first phase (***Analysis of CMMI practices***), we analyzed CMMI-DEV documentation (software engineering process area/specific practices) searching for any citation of HCI issues. We found some quotes relating to HCI issues for the majority of practices (27), except for: two practices of RD (SP2.2 and SP2.3); three practices of TS (SP2.2, SP2.3, and SP2.4); and almost all practices of PI (we found citations only for SP1.1).

The practices of RD are more related to functional aspects and not to HCI concerns. Technical Solution practices are related to the functional software components (technical data package, reuse of product component designs, etc.). Finally, the PI process area is responsible for achieving complete product integration through the progressive assembly of product components (i.e. service, service systems, and their components). We found citations only regarding strategy definition to perform the product integration.

After finding all citations, we organized them separately to identify the main approaches related to HCI. Figure 3 presents, as an example, the HCI categories that support specific practices of Verification and Validation process area.

We note that one HCI category (composed of HCI methods, techniques, standards, patterns, and other approaches) can support different practices of the same/different CMMI-DEV process area (composed of engineering best practices). Moreover, one specific CMMI-DEV practice can be supported by several HCI categories (for instance, VAL SP1.1 is supported by *Evaluation methods for HCI verification tests* and *Functional Prototype to validate HCI*).

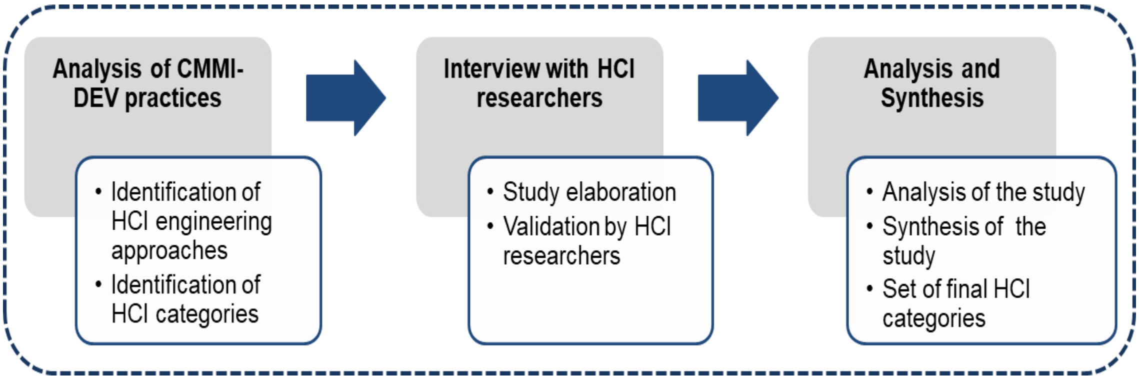

Figure 2. Research Methodology (adapted from (Gonçalves et al., 2016))

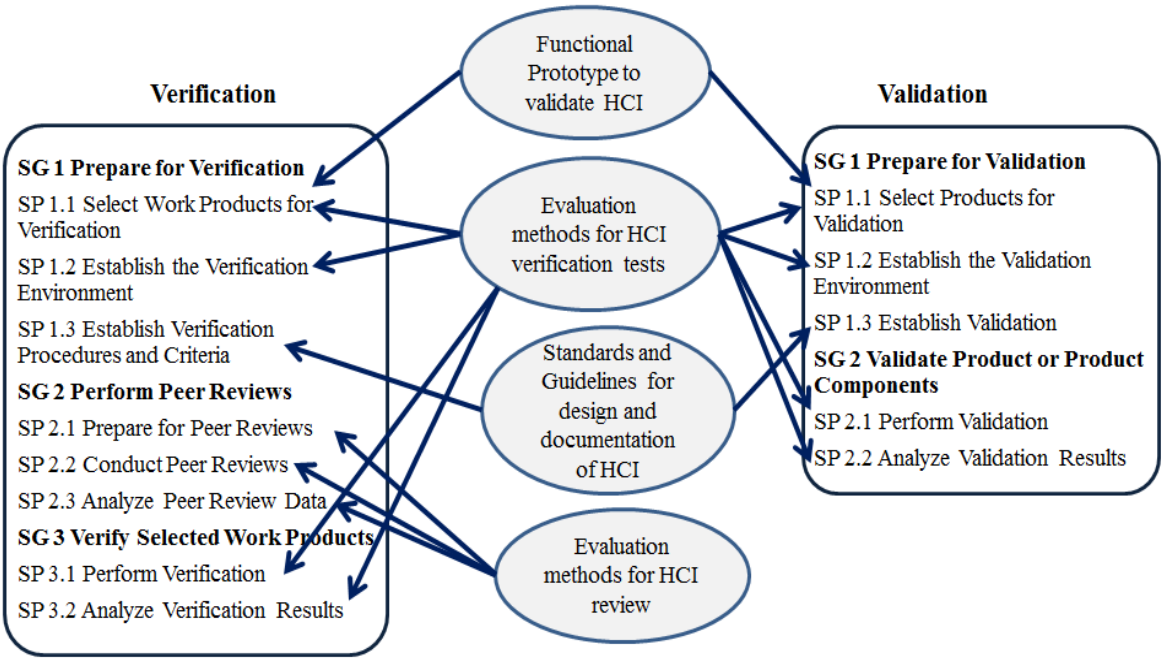

Figure 3. Verification and Validation specific practices (CMMI-DEV) x HCI categories

We identified ten HCI categories, as follows:

1) Task Analysis Methods for HCI
2) Prototype for HCI requirements
3) Operational Concepts and Scenarios Specification for HCI
4) Standards and Guidelines for design and documentation of HCI
5) Techniques to validate HCI requirements
6) Architecture Patterns for HCI
7) Design patterns for HCI
8) Functional Prototype to validate HCI
9) Evaluation methods for HCI verification tests
10) Evaluation methods for HCI review

In the second phase (***Interview with HCI researchers***), we planned and performed an experimental study ***interviewing HCI researchers*** in the HCI domain. The objective was to evaluate and improve the propositions of HCI categories and examples (Gonçalves et al., 2018). Our planning considered the following protocol information:

- Object of the study: HCI categories (with examples of HCI approaches) and engineering practices of the CMMI-DEV model (see Table 2).
- Objective: To validate and improve the proposed HCI categories and examples to support the engineering practices of the CMMI-DEV.
- Subjects: Respondents from HCI domain with a Ph.D. degree from different countries (see Table 1). All respondents are professors in universities and some of them (13/20) have industry experience in the HCI domain.

Table 1. Information about HCI researchers (adapted from (Gonçalves et al., 2018))

CS = Computer Science, SE = Software Engineering and HCI = Human-Computer Interaction.

HCI researcher	Background				Origin	Interview	
	Years of experience	Industry Experience	Ph.D. domain	Current interest in interactive systems		Duration	Type
1	13	No	HCI	Methods and models for HCI design and evaluation	France	01h30	In person
2	25	Yes	HCI	Tools for design, realization and evaluation	France	00h55	In person
3	8	No	HCI	Agent-based architecture models and HCI evaluation	France	01h00	In person
4	8	Yes	SE-HCI	Interaction and Automatic Reasoning	France	00h50	In person
5	25	Yes	SE-HCI	Methods and tools of systems engineering	France	01h15	In person
6	26	No	HCI	HCI	France	00h50	In person
7	27	Yes	SE-HCI	SE and HCI	Belgium	00h50	In person
8	20	Yes	HCI	HCI	Brazil	02h17	Video conference
9	10	No	HCI	HCI	Brazil	00h40	Video conference
10	25	No	HCI	HCI	France	01h00	In person
11	20	Yes	SE-HCI	User Interfaces Plasticity, Creativity Support Tools, and Persuasive Technology	France	01h45	In person
12	40	Yes	SE-HCI	Innovative interfaces, mobility	France	01h30	In person
13	12	Yes	SE-HCI	Quality of Human-Computer Interfaces	France	00h53	In person
14	7	Yes	SE-HCI	HCI	France	01h00	In person
15	10	No	HCI	HCI	Brazil	01h03	Video conference
16	30	Yes	CS-HCI	Interactive critical systems	France	01h36	Video conference
17	27	Yes	CS-HCI	HCI design, Ubiquitous computing	Tunisia	01h26	Video conference
18	21	Yes	CS-HCI	Semiotic engineering, evaluation, and design of interfaces	Brazil	01h39	Video conference
19	10	Yes	CS-HCI	Organizational Semiotics, Culture and Values in design	Brazil	01h03	Video conference
20	27	Yes	CS-HCI	Service Design, Ubiquitous Computing, SOA	Algeria	01h50	In person

<table>
<tr><th rowspan="2">Process Area and Specific Goal (SG)</th><th rowspan="2">Specific Practice (SP)</th><th rowspan="2">Methods, techniques, standards, and patterns of HCI</th><th colspan="3">Answer</th><th rowspan="2">Justification</th></tr>
<tr><th>I agree</th><th>I partially agree</th><th>I do not agree</th></tr>
<tr><td rowspan="2">Verification

SG 1 Prepare for Verification

Preparation for verification is conducted.</td><td rowspan="2">SP 1.1 Select Work Products for Verification

Select work products to be verified and verification methods to be used.</td><td>Evaluation methods for HCI verification tests</td><td rowspan="2"></td><td rowspan="2">X</td><td rowspan="2"></td><td rowspan="2">“To add other types of tests. To use classical tests of verification.”</td></tr>
<tr><td>Examples:
• Usability tests
☐ Exploratory tests
☐ Assessment tests
☐ Validation or verification tests
☐ Comparison tests
• Validation by experts</td></tr>
</table>

Figure 4. Extract of the questionnaire

They were selected by convenience from the research contacts of one of the authors, and/or considering their reputation in international HCI conferences and community.

- Instrumentation: A specific questionnaire composed of 33 items and elaborated in three languages: English, French, and Portuguese. Each item corresponds to a proposition of one HCI category to one specific practice (see Table 2). Figure 4 presents part of the questionnaire related to specific practices of Verification and Validation process areas.
- Experimentation: Respondents were interviewed in person or by video conference (see Table 1). Before starting the interview, we presented and explained the process areas of the CMMI-DEV to each respondent. One/two researchers performed the interviews. For each practice, the respondents answer if they agree (*A*), partially agree (*a*) or do not agree (*D*) that the associated HCI category supports the practice. The partially agree and disagree responses were justified by the respondents, and when necessary they explained other proposals.

- Analysis of the results: Since all interviewees were researchers in HCI the domain (from 7 to 40 years of experience), all the suggestions proposed by the respondents were analyzed in the same way.

Following this protocol, we interviewed 20 HCI researchers from five different countries (12 respondents from France, 5 from Brazil, 1 from Algeria, 1 from Tunisia and 1 from Belgium).

Considering the rather low size of our participant sample (20) but with many respondents from France (12), we decided to set only two main origins for this study: France and other.

Finally, in the last phase ***analysis and synthesis***, we performed a qualitative analysis to have a final proposal of HCI categories for each practice. For this, we organized all comments written by each HCI researcher for each evaluated item. Three authors of this paper participated in this analysis and synthesis (see Figure 5).

First, one of the authors (here named author 1) analyzed the comments to identify if there was a proposal categories improvement, or new approaches. The comments were organized in groups that present similar proposals.

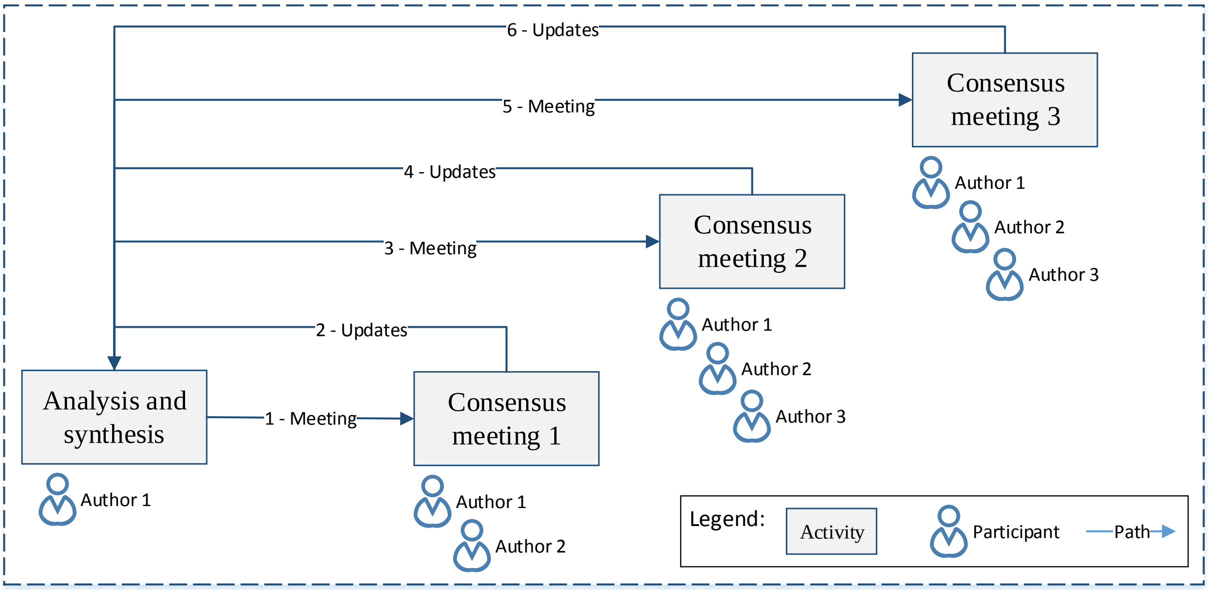

Figure 5. Analysis and synthesis process

Table 2. Items (variables) of the study (for each row, the bold letters of the three columns are used to build the variable label, indicated in column 4)

Process Area (PA) and Specific Goal (SG) (CMMI Product Team, 2010)	**Specific Practice (SP)** (CMMI Product Team, 2010)	**HCI categories (methods, techniques, standards, and patterns)**	**Variables initials**
Requirements Development *SG 1 Develop Customer Requirements*	SP 1.1 **E**licit Needs	Task Analysis **M**ethods for HCI **P**rototype for HCI requirements	REM REP
Requirements Development *SG 1 Develop Customer Requirements*	SP 1.2 **T**ransform Stakeholder Needs into Customer Requirements	Task Analysis **M**ethods for HCI	RTM
Requirements Development *SG 2 Develop Product Requirements*	SP 2.1 Establish **P**roduct and Product Component Requirement	Task Analysis **M**ethods for HCI	RPM
Requirements Development *SG 3 Analyze and Validate Requirements*	SP 3.1 Establish **O**perational Concepts and Scenarios	**O**perational Concepts and Scenarios Specification for HCI	ROO
	SP 3.2 Establish a **D**efinition of Required Functionality and Quality Attributes	**S**tandards and Guidelines for design and documentation of HCI	RDS
Requirements Development *SG 3 Analyze and Validate Requirements*	SP 3.3 **A**nalyze Requirements	Task Analysis **M**ethods for HCI	RAM
	SP 3.4 Analyze **R**equirements to Achieve Balance	**T**echniques to validate HCI requirements	RRT
	SP 3.5 **V**alidate Requirements	**P**rototype for HCI requirements	RVP
Technical Solution *SG 1 Select Product Component Solutions*	SP 1.1 **D**evelop Alternative Solutions and Selection Criteria	**A**rchitecture Patterns for HCI	TDA
	SP 1.2 Select Product **C**omponent Solutions	**O**perational Concepts and Scenarios Specification for HCI	TCO
Technical Solution *SG 2 Develop the Design*	SP 2.1 Design the **P**roduct or Product Component	**P**rototype for HCI requirements **A**rchitecture Patterns for HCI **S**tandards and Guidelines for design and documentation of HCI	TPP TPA TPS
Technical Solution *SG 3 Implement the Product Design*	SP 3.1 **I**mplement the Design	Design patterns for HCI	TID
Technical Solution *SG 3 Implement the Product Design*	SP 3.2 Develop Product **S**upport Documentation	**S**tandards and Guidelines for design and documentation of HCI	TSS
Product Integration *SG 1 Prepare for Product Integration*	SP 1.1 **E**stablish an Integration Strategy	**P**rototype for HCI requirements **F**unctional Prototype to validate	PEP PEF

Process Area (PA) and Specific Goal (SG) (CMMI Product Team, 2010)	**Specific Practice (SP)** (CMMI Product Team, 2010)	**HCI categories (methods, techniques, standards, and patterns)**	**Variables initials**
		HCI	
V**a**lidation *SG 1 Prepare for Validation*	SP 1.1 **S**elect Products for Validation	Evaluation methods for HCI **v**erification tests	ASV
		Functional Prototype to validate HCI	ASF
	SP 1.2 **E**stablish the Validation Environment	Evaluation methods for HCI **v**erification tests	AEV
	SP 1.3 Establish **V**alidation Procedures and Criteria	**S**tandards and Guidelines for design and documentation of HCI	AVS
V**a**lidation *SG 2 Validate Product or Product Components*	SP 2.1 **P**erform Validation	Evaluation methods for HCI **v**erification tests	APV
	SP 2.2 **A**nalyze Validation Results	Evaluation methods for HCI **v**erification tests	AAV
V**e**rification *SG 1 Prepare for Verification*	SP 1.1 **S**elect Work Products for Verification	Evaluation methods for HCI **v**erification tests	ESV
		Functional Prototype to validate HCI	ESF
	SP 1.2 **E**stablish the Verification Environment	Evaluation methods for HCI **v**erification tests	EEV
	SP 1.3 Establish **V**erification Procedures and Criteria	**S**tandards and Guidelines for design and documentation of HCI	EVS
V**e**rification *SG 2 Perform Peer Reviews*	SP 2.1 Prepare for Peer **R**eviews	Evaluation methods for HCI **r**eview	ERR
	SP 2.2 **C**onduct Peer Reviews	Evaluation methods for HCI **r**eview	ECR
	SP 2.3 Analyze Peer Review **D**ata	Evaluation methods for HCI **r**eview	EDR
V**e**rification *SG 3 Verify Selected Work Products*	SP 3.1 **P**erform Verification	Evaluation methods for HCI **v**erification tests	EPV
	SP 3.2 **A**nalyze Verification Results	Evaluation methods for HCI **v**erification tests	EAV

Then a second author reviewed the proposal and proposed some modifications. Authors 1 and 2 reached a consensus in a meeting and updated the analysis and synthesis. Then, the third author (the most experienced in HCI domain) reviewed the change and a discussion meeting took place. The three authors reached a consensus and a new update of the analysis and synthesis document was made. Finally, a third consensus meeting 3 with the three authors was performed for a last review of the proposals (this meeting took three hours). A detailed description of this first study can be found in (Gonçalves et al., 2018).

4. Investigating agreement among HCI researchers

While analyzing the results of our first study (the third phase - see Figure 2), we identified that some proposals had more statistical agreement among the researchers than others.

We decided, therefore, to perform a statistical analysis of the HCI researchers' agreements. In this section, we describe the detailed information (instrument and the items being evaluated) that we used in the statistical analysis.

4.1. Detail of the Instrument and the Field Study Data

Table 2 presents the 33 items that were evaluated and compose the instrument. As previously described, a specific questionnaire was elaborated in three languages: English, French, and Portuguese. Before the elaboration of the questionnaire, we performed the ***Analysis of CMMI practices*** (see Figure 2 and section 3).

To facilitate the statistical analysis, we coded each item (Table 2) with three capital letters. The first one is the letter of the corresponding process area name, the second and the third ones are the initial letter of one of the names that compose the specific practice and the approach category, respectively (marked in bold in Table 2). Each item code is presented in the last column of Table 2. There are also four variables related to the HCI researcher's main characteristics (years of experience, industry experience, Ph.D. domain, and origin).

4.2. Data analysis

Even though the number of individuals is low (20), the presence of a rather large number of variables (37 = 33 investigated specific practices + 4 HCI researcher's main characteristics) shows that a descriptive multivariate statistical approach is preferable to a monovariate one. The main study output data set can be seen as a table with two entries where the rows correspond to the **20 HCI researchers** and the columns to the variables, namely the **33 variables** linked to the questionnaire (Table 2) and the **4 variables** related to HCI researcher's main characteristics.

When the majority of variables have a qualitative measurement scale model, Multiple Correspondence Analysis – MCA can be used (Benzécri, 1992) (Nishisato, 2007) (Beh & Lombardo, 2014) (Friendly & Meyer, 2016). To get homogeneous data, the variable with a quantitative measurement scale, namely *Experience* (between 7 – named *HCI researcher*, and 40 years – named *Senior HCI researcher*) was changed into a qualitative one using fuzzy coding (to reduce the information loss and to avoid the delicate choice of thresholds between adjacent modalities).

Given the low number of individuals, only two modalities are considered, Figure 6 (Benzécri, 1992). MCA being less known than usual multivariate methods for quantitative variables (e.g. Principal Component Analysis (PCA), hierarchical clustering or multiple regression), two aspects are worth reviewing:

- MCA is based on the same principle of PCA (Johnson & Wichern, 1992) but deals with qualitative variables. That means each data piece is mainly within the set {0, 1} according to such modality with such variable is either absent or present for a given individual. Considering the 33 variables linked to the questionnaire, this coding approach is interesting when the HCI researcher does not provide any answer because a new modality corresponding to ***N**o answer* can be added to the three initial ones, i.e. ***A**gree*, *partially **a**gree* and ***D**isagree*. These 4 modalities will be labeled *N*, *A*, *a*, and *D* respectively. The binary sub-table corresponding to an HCI researcher giving *No answer* for a given variable is coded as (1, 0, 0, 0). Still considering these 33 variables, the number of modalities runs between 2 and 4, e.g. 2 corresponds to the case where only *a* and *A* are used (for ten questions). Given all modality possibilities for the 37 variables, the total number of modalities is 100, thus yielding a membership value table with a rather large number of columns (it is worth reminding that values within the interval [0, 1] are obtained with one variable, see Figure 6);

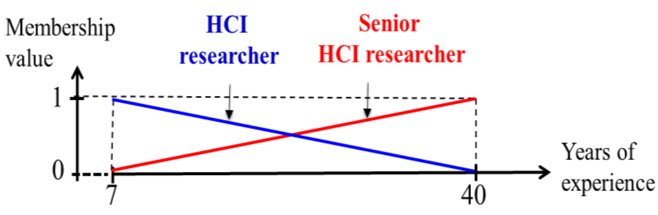

Figure 6. Fuzzy coding of the variable *number of years of experience of the HCI researchers*

- as with PCA, MCA may consider input statistical objects (rows or columns) with either an active or passive status. This possibility is interesting in our case (see Figure 7). The 33 variables linked to the questionnaire will have an active status (it is important to find the relationships between the HCI researchers opinions); and the 4 variables related to HCI researchers characteristics will have a passive status (these variables may give some information to explain the previous relationships and inter-individual differences; the corresponding modality points are often named *supplementary points*) (Benzécri, 1992).

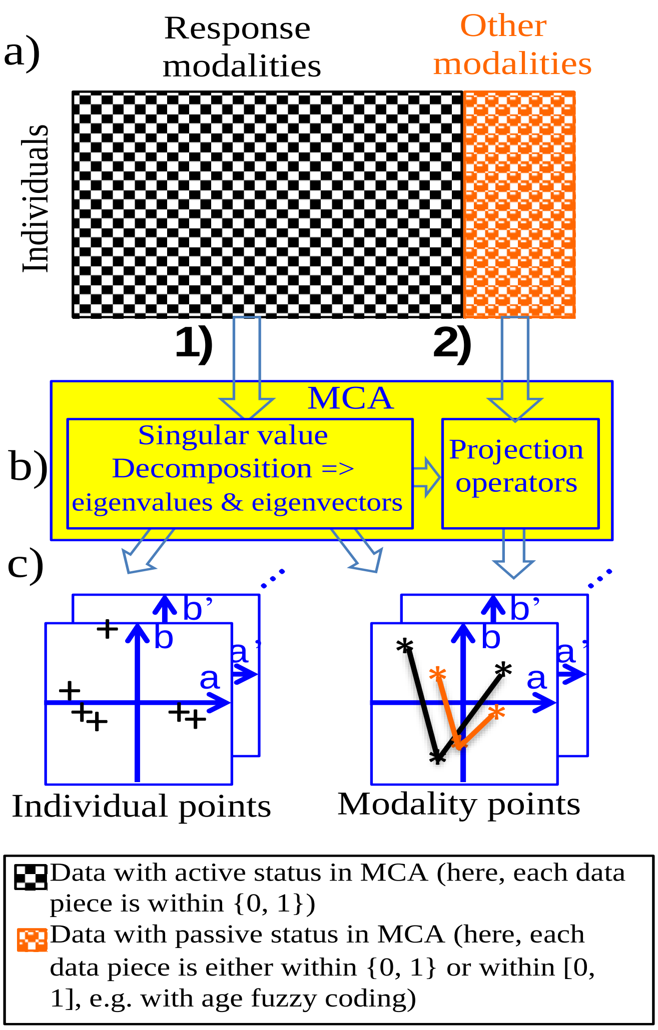

Figure 7. The overall principle of Multiple Correspondence Analysis (MCA) with our data. (a) Input table (*other modalities* correspond to 4 variables related to HCI researcher's characteristics, with one variable being present through 2 fuzzy modalities, Figure 6), (b) MCA with its possibility to consider supplementary column points, (c) some MCA outputs (other outputs are eigenvalues in % and small tables that aid interpretation).

The statistical analysis was performed using the statistical software *R* with the package *FactoMineR* (Cornillon, 2018). This package yields MCA outputs with multiple aspects such as graphics, small tables that aid interpretation (e.g. point contributions, in %, in the positioning of a main axis) or sorting procedure outputs concerning the hypothesis test of the link between a main axis and the 37 variables or the 100 modalities (inference statistics).

Once the main results were found from MCA, more conventional techniques were used for graphical views or hypothesis tests such as the connection between two variables. The number of HCI researchers being low, there is a possibility to get expected frequencies less than 4 or 5, Pearson's chi-squared test is not recommended, even with the Yates's correction, see (Sheskin, 2007) for discussion about such limitations. That is why Fisher's exact test for count data was used (Crawley, 2007). For the statistical tests, the usual 5% confidence level is used. These tests are present in *R*, but some post hoc tests and graphical views require using specific packages (*RVAIdeMemoire* and *Lattice)*.

5. Analysis and Discussion

In this section, we will present the analysis and discussion of the data following monovariate and multivariate approaches.

5.1. Monovariate Analysis

A preliminary monovariate analysis, namely studying the way the answer scale (*Agree*, *partially agree*, *Disagree*, *No Answer*) is used.

Figure 8 (a) shows overall results for the five process areas:

- Except for Requirements Development (R), all process areas have more than 50% agreement;
- Validation (A) and Product Integration (P) are the process areas with least (1%) and most (27%) disagreements respectively;
- Requirements Development (R) has almost the same number of agreements and partially agreements (47% and 45% respectively);
- Verification (E) is the process area with most agreements (68%); and,
- Although Technical Solution (T) has the widest range of agreement (50%), it has two items not evaluated (by two respondents).

Figure 8 (b) and (c) show relevant results for the 33 items and 20 HCI researchers. We can note that:

- the *Agree* modality is the most often (384 times – 58%) used one. It is worth noting that given both the chi-square distance and geometrical principles of MCA, the differences in the way the modalities are used will be displayed.
- TPA, ERR, and ECR are the variables with the most occurrence of *Agree* modality (85%).
- RTM, RPM, TPS, ASV, AEV, AAV, AVS, EVS, ERR, ECR, and EDR did not present *Disagree*, and *No answer* modalities, and for 9 of these variables the modality most often used was *Agree*.
- all variables present *Agree* and *partially agree* modalities.
- 21 variables present *Disagree* modality, being PEP the variable with the most occurrences (45%); 4 variables (RRT, TID, APV, and EEV) present *No answer* modality, which was used by three HCI researchers (*HCI researchers* 4, 10 and 18).
- Finally, there are considerable differences, considering the **HCI researcher** factor. For instance, HCI researcher 2 always *Agreed* while HCI researcher 18 often *partially agreed*; HCI researcher 3 often *Disagreed*.

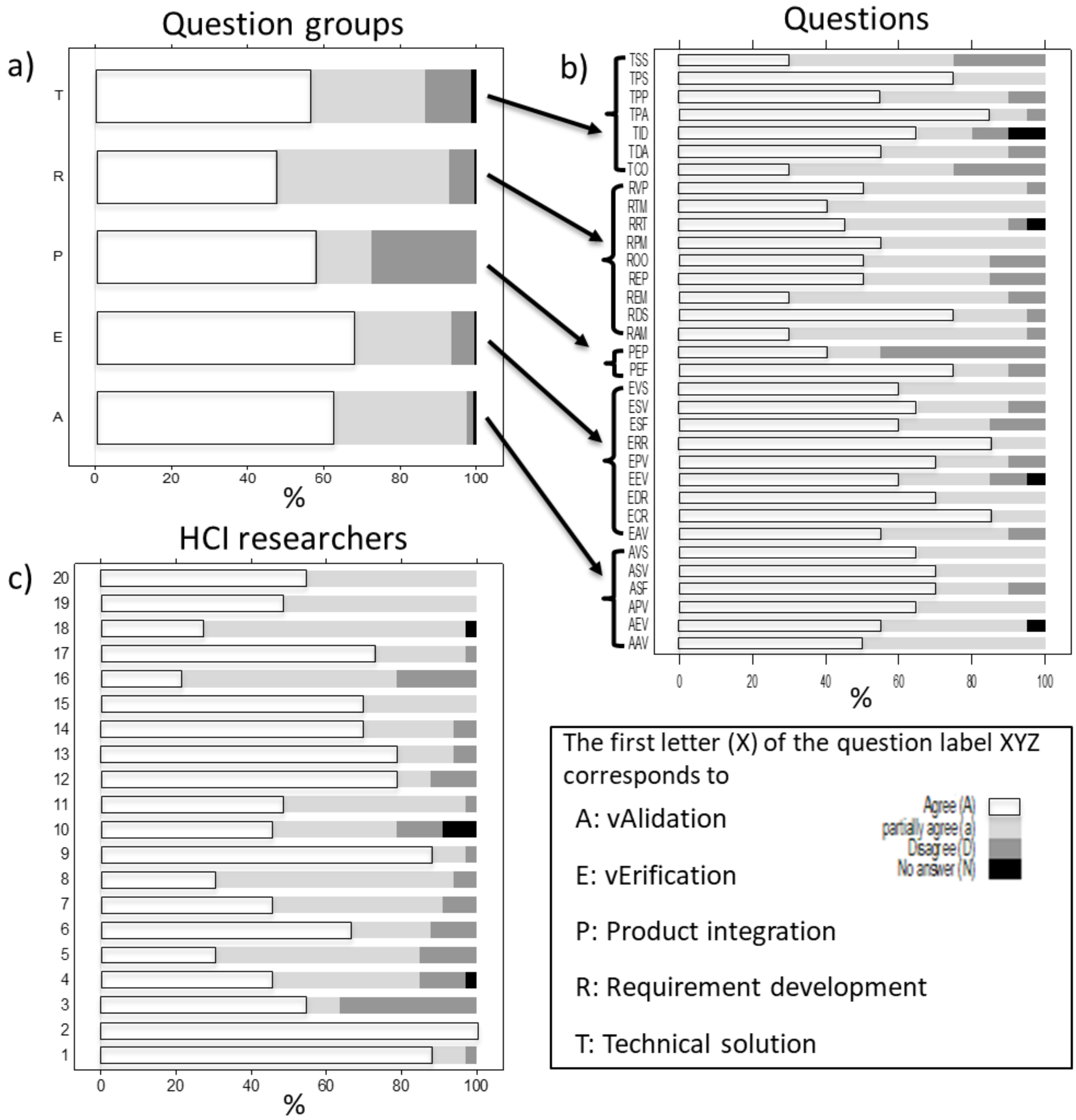

Figure 8. Use of the four modalities for (a) the 5 groups of questions, (b) the 33 questions and (c) the 20 HCI researchers

5.2. Monovariate results discussion

Considering the monovariate results, Figure 8 (a) and (b) shows that variables such as ERR, ECR, and EDR had an excellent agreement among HCI researchers because they are part of the same process area (Verification - see Table 2). Also, they are used together to achieve the same goal (SG 2 Perform Peer Reviews - see Table 2). On the other hand, variables such as TCO and TSS had little agreement among HCI researchers, because they are used to achieve different goals even if they are part of the same process area (Technical Solution - see Table 2).

Figure 8 (c) shows that some evaluations do not present disagreement (*HCI researchers 2, 15, 18, 19 and 20*) or an expected number of disagreements with respect to size of the domain: 1 disagreement (*HCI researchers 1, 9, 11 and 17*); 2 disagreements (*HCI researchers 8, 13 and 14*); 3 disagreements (*HCI researcher 7*); and 4 disagreements (*HCI* *researchers 4, 6, 10 and 12*). These disagreements support to discuss the illustrations of techniques and methods and update of the initial proposals, as well as the categories name and, sometimes, the inclusion of new categories. A smaller number of HCI researchers express most disagreements as follows:

- 5 disagreements (for *HCI researcher 5*) and 7 disagreements (for *HCI researcher 16*) - these respondents are HCI researchers in the design and evaluation of critical interactive systems, also supporting modifications in initial proposals or examples;
- 12 disagreements (for *HCI researcher 3*) - the interactions with this HCI researcher show the willingness to deepen the propositions by questioning him, to provoke a discussion; and in the end, the HCI researcher had few suggestions of improvement.

The diversity of opinions to *partially agree* is also noteworthy: from 0 (*HCI researcher* 2) to 23 times (*HCI researcher* 18) without *disagreements*. But this is not surprising because the respondents have experience in different domains. They contributed to refinements of the initial proposals. Histograms related to variables (Figure 8 (a) and (b)) are often built, and histograms related to individuals (Figure 8 (c)) are least often considered. The latter can be built because each of the 33 variables has the same scale mathematical model, i.e. nominal qualitative with 4 modalities and summing counts across the 33 variables may yield information about the individual nature.

In our case, one can say that some individuals may have a most "protesting" character as they have a stronger tendency to contradict the proposals submitted to them. This is particularly the case of *HCI researcher 3.*

On the other hand, one can find a "conformist" trait such as *HCI researcher 2* who uses *Agree* modality in 100% cases. This HCI researcher is, in fact, very familiar with HCI approaches regarding the categories discussed. Over 25 years of experience in HCI, he often had the opportunity to participate in projects and thesis supervision with two of the co-authors. Knowledge of the HCI field was much frequently shared in this case.

Between the two different character traits, there is a "hesitant" e.g. with 70% (*HCI researcher 18*) of *partially agree* modality. The responses *partially agree* of this respondent contributed to the refinements of the initial proposals and sometimes in the inclusion of new categories in the first study (Gonçalves et al., 2018).

5.3. Multivariate Analysis

The percentage sequence of variance explained by the main axes is 18%, 13, 12, 9, …, which corresponds to rather high values with so many modalities. Let us focus on the main plane that crosses Principal axes 1 and 2, i.e. 31% of inertia. Axis 1 is mainly controlled by modality *a* for 11 questions; see the right side of Figure 9, for the variables ESV, EEV, EPV, ERR, ECR and AEV that have the 6 highest contributions.

The relative positions of *A* and *a* modality points along Axis 1 show that: (1) Axis 1 mainly opposes modality *A*, on the left side, to modality *a*, on the right side; and (2) the former is more often used than the latter (according to the usual barycenter principle in mechanics, *A* modality points are closer to the center of gravity, i.e. the intersection between the main axes).

Figure 10 is consistent with this result since it opposes a cluster of 9 HCI researchers on the left side (thus often using modality *A* for questions linked to Axis 1, see Figure 9) to a cluster of 5 HCI researchers on the right side (using modality *a*).

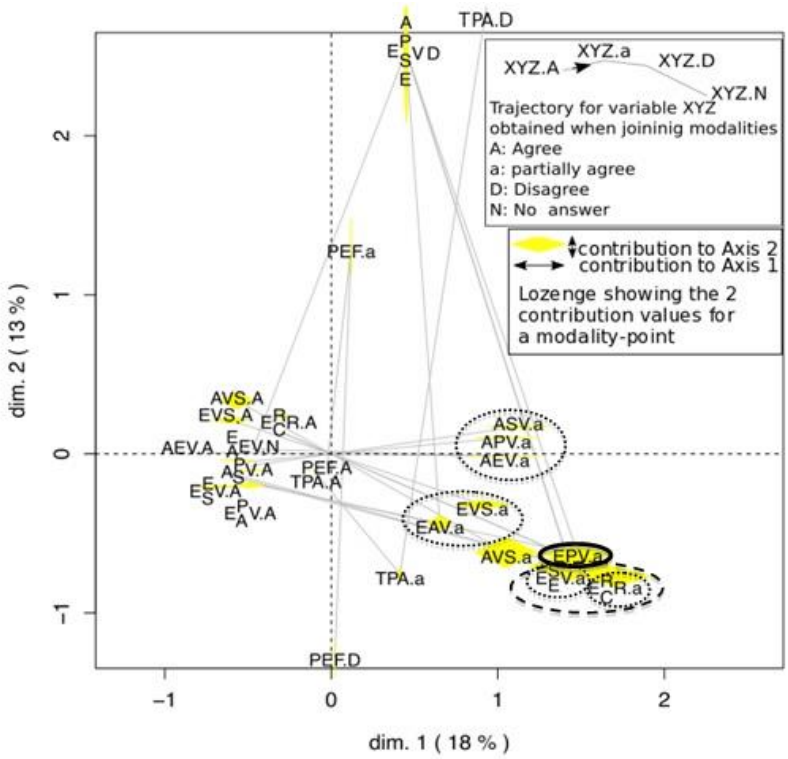

Figure 9. Main plane from MCA of a table crossing the 20 HCI researchers and the 91 modalities of the 33 questions. Projection of the 91 modalities on the main plane. Only the questions having modalities with high contributions to control Axes 1 and 2 are shown; for each question, the modalities are linked with the following sequence: *A*, *a*, *D*, *N*. If a modality does not exist (no HCI researcher uses this modality), the corresponding point is removed.

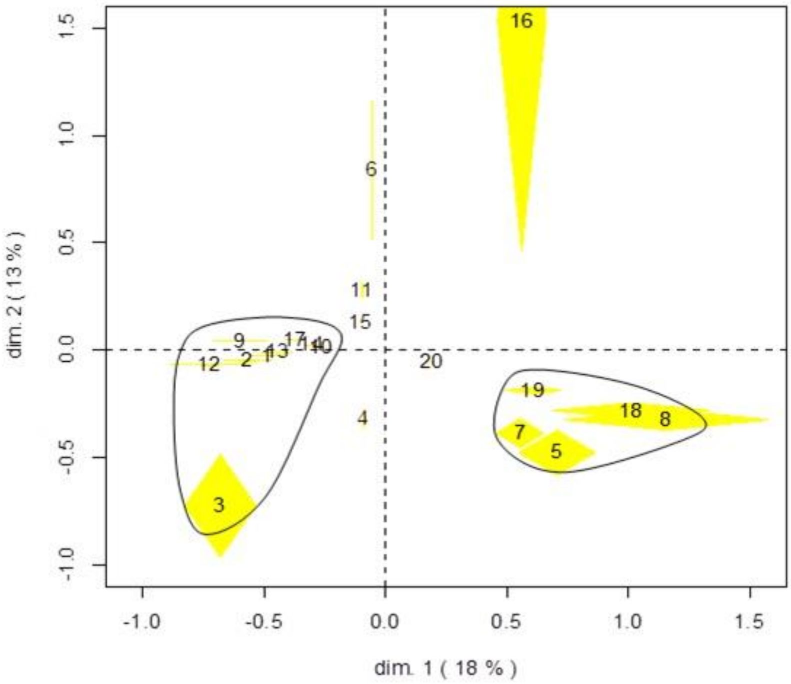

Figure 10. Main plane from MCA of a table crossing the 20 HCI researchers and the 91 modalities of the 33 questions. Projection of the 20 HCI researchers: on the left side, a cluster of 9 HCI researchers mainly using modality *A* for the questions well linked to Axis 1, see Figure 9; on the right side, idem but with a cluster of 5 HCI researchers often using modality *a*.

Table 3, which describes the responses for the two clusters of HCI researchers, confirms the trends shown by Axis 1.

For instance, on the very right side of Axis 1 and bottom of Table 3, *HCI researcher 8* uses modality *a* for each of the 11 variables, while *HCI researcher 7* uses modality *a* 8 times out of 11. On the other side of Axis 1, the same distinction could be performed for *HCI researcher 3* vs. *HCI researcher 10*.

Observing Figure 9, Figure 10, and Table 3, the following three remarks must be made:

1. Being in a multivariate statistical context, a principal axis shows the main trends. For instance, the 9 individuals situated on Axis 1 left side present a strong propensity to *agree* with 11 proposals; this does mean that each of the 9 HCI researchers uses modality *A* for each of the 11 variables (otherwise, the 11 individuals would have the same place along Axis 1).
2. Being in the factor analysis general context, MCA (as PCA) exhibits phenomena hierarchically. For instance, Axis 1 shows stronger relational phenomena than Axis 2, namely correspondences between the modalities having the main role in Axis 1 control are more obvious than correspondences shown by Axis 2. In our case, the correspondences displayed by Axis 1 are present for 11 variables and 14 individuals, while both the numbers of variables and individuals are lower for Axis 2, see Tables 3 and 4.
3. Being in the MCA context (that means, not in PCA context), the variables control the principal axes through their corresponding modalities. For Axis 1, the control level, which is shown using a lozenge size according to the horizontal direction, is higher for modality *a* than *A*; that is why the 11 variable labels shown in Table 3 left side are displayed with modality *a*.

Table 3. Main results underscored by Principal Axis 1. The columns correspond to the most discriminant and connected modalities; the rows to 2 clusters of HCI researchers (the 14 HCI researchers present significant contributions in Axis 1 control); a row/column intersection gives the HCI researcher's response with *A*=Agree, *a*=partially agree and *N*=No answer.

		Variable/Modality pairs playing a main role in Principal Axis 1 control. The contributions of the 11 pairs are displayed in decreasing order (from left to right)										
	Variable	ESV	EEV	EPV	ERR	ECR	AEV	AVS	APV	ASV	EVS	EDR
	Modality	a	a	a	a	a	a	a	a	a	a	a
Cluster of HCI researchers situated on Axis 1 left side	3	A	A	A	A	A	A	A	A	A	A	A
	12	A	A	A	A	A	A	A	A	A	A	A
	9	A	A	A	A	A	A	A	A	A	A	A
	2	A	A	A	A	A	A	A	A	A	A	A
	1	A	A	A	A	A	A	A	A	A	A	A
	13	A	A	A	A	A	A	A	A	A	A	A
	17	A	A	A	A	A	A	A	A	A	A	A
	14	A	A	A	A	A	A	A	A	A	A	A
	10	A	N	A	A	A	N	A	A	A	A	a
Cluster of HCI researchers situated on Axis 1 right side	7	a	a	a	A	A	a	a	A	a	a	a
	19	a	a	a	A	A	a	a	a	a	a	A
	5	a	a	A	a	a	a	a	a	A	a	a
	18	a	a	a	a	a	a	a	a	a	a	a
	8	a	a	a	a	a	a	a	a	a	a	a

Table 4. Main results underscored by Principal Axis 2. The columns correspond to the most discriminant and connected modalities; the rows to 2 clusters of HCI researchers (the 3 HCI researchers present high contributions in Axis 2 control); a row/column intersection gives the HCI researcher's response with *A*=Agree, *a*=partially agree, *D*=Disagree and *N*=No answer.

	Variable/Modality pairs playing a main role in Principal Axis 2 control. The contributions of the 6 pairs are displayed in decreasing order (from left to right)						
	Variable	ESV	EEV	EPV	EAV	TPA	PEF
	Modality	D	D	D	D	D	a
HCI researchers situated on Axis 2 top side	16	D	D	D	D	D	a
	6	D	D	D	D	A	A
HCI researcher situated on Axis 2 bottom side	3	A	A	A	A	A	D

These three remarks must be kept in mind for the rest of the analysis. Given the results drawn from points with an active case in MCA, i.e.:

- Relationships between the modalities, Figure 9;
- Resemblances between the HCI researchers, Figure 10.

Correspondences between the modalities and the HCI researchers (Table 4) being found; the focus on the 4 variables related to HCI researcher's characteristics, namely on 9 modalities, could be helpful to explain these correspondences.

Figure 11 shows that the variables related to HCI researcher's characteristics have little ability to explain these correspondences; except maybe for **Origin** variable, e.g. the right side position of *Other* modality (Figure 11) means that this HCI researcher category more often uses the modality *a* (see Figure 9 right side), while *French* HCI researchers more often used modality *A* (Figure 9 left side).

In most cases, users of descriptive methods such as PCA or MCA do not use the inferential possibilities of these methods.

The package FactoMineR includes a function (*dimdesc*) that aids interpretation concerning the variability of each main axis (Cornillon, 2018).

This function sorts the variables and the modalities (both with active and passive case) using a One-way ANOVA as follows: the variable *Y* corresponds to one main axis (e.g. coordinates of the 20 individuals according to Axis 1) and is explained according to a qualitative variable (e.g. variable *X* is the **Origin**).

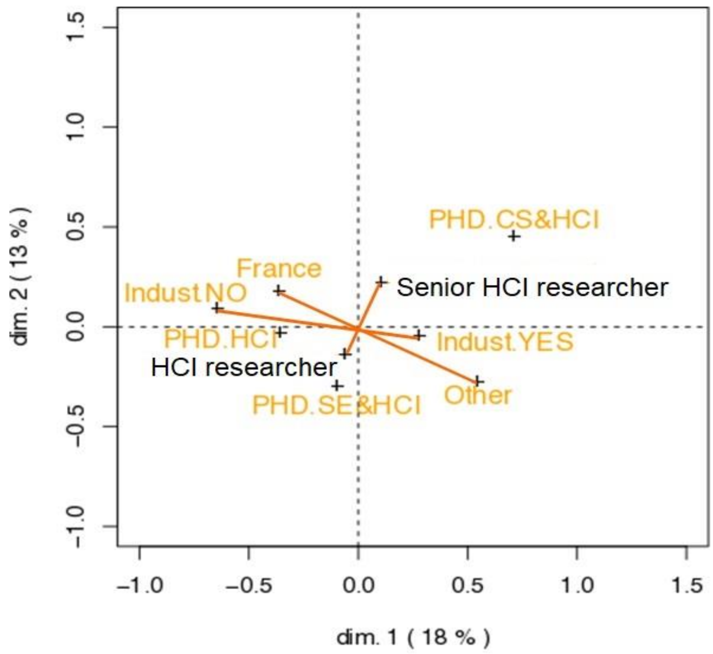

Figure 11. Main plane from MCA of a table crossing the 20 HCI researchers and the 91 modalities of the 33 questions. Projection of additional columns points corresponding to the 4 variables related to HCI researcher's characteristics (see stage (2) of the Figure 7 for the overall principle).

For Axis 1, the testing procedure shows p-values (indicated below in parentheses) sorted as follows: (1) ESV (2 10^{-6}), (2) EEV (1 10^{-5}), (3) AEV (2 10^{-5}), … , (16) *Origin* (5 10^{-2}), … (in fact, all the variables indicated in Table 3 shows a p-value lower than 10^{-3}).

These p-values, which are consistent with Figures 8 and 9, confirm that the explanatory power of HCI researcher's characteristics is poor, and the highest level corresponds to the **Origin** variable.

Fisher's exact test, performed between this variable and the main variables according to Axis 1, shows only 3 cases where the Null hypothesis[1] must be rejected, i.e. with EPV, AEV, and ASV. The same test performed with **Industry** variable yields that the Null hypothesis must always be accepted, which is consistent with MCA.

Axis 2 is mainly positioned by modality *D* of variables ESV, EEV, EPV, EAV, and TPA, see Figure 9 top. These close positions are mainly due to *HCI researcher 16*, Figure 10 top, who uses the *Disagree* modality for these 5 questions. *HCI researcher 6* and *HCI researcher 3* also have the main contribution in Axis 2 control. *HCI researcher 6* uses modality *D* for ESV, EEV, EPV, EAV but not TPA, yielding a moderate top position compared with the extreme top position of *HCI researcher 16*. The opposite position of *HCI researcher 3* is due to the absence of use of modality *D* for these 5 variables. As in Axis 1, the 4 variables related to HCI researchers' characteristics present little ability to explain the difference between these 3 HCI researchers. Table 4 displays the main trends underlined by Axis 2.

5.4. Multivariate result discussion

Given the possibility of relationships between the **37 variables** (the 33 questions plus the 4 variables linked to HCI researchers), the following preliminary remarks must be made: using a multivariate method once is more complex than using 37 times a monovariate method (or many times a bivariate method). Then, if a multivariate method as MCA is descriptive, this

[1] We do not have any Null hypothesis about our study. In this paper, Null hypothesis refer to statistical tests being more or less know (Fisher's exact test to show that two variables are independent, (parametric) ANOVA test with MCA and so on).

one does not require probabilistic hypotheses (e.g. the presence of *a priori* specific distribution models). Keeping in mind these remarks, our main objective was to draw taxonomic dimensions between our 20 HCI researchers, even though we are aware that the generalization of MCA results is very difficult.

For the multivariate approach, MCA was used even though this one is both little-known and more complex than other techniques, e.g. Principal Component Analysis or Hierarchical Clustering. Analyzing Figure 9, we note that some much-closed item is semantically associated as follows (see different circles):

- ESV and EEV (dotted circle) correspond to the preparation for applying verification techniques (identification of products to be verified and preparation of environment) - *(1)*; ERR and ECR (dotted circle) correspond to prepare and execute the peer review (one category of technique) - *(2)*;
- *(1)* and *(2)* (put together using a dashed circle) are correlated since *(1)* are defined to be used in *(2)*;
- *(1)* and EPV (continued line circle) are correlated since *(1)* are defined to be used in EPV;
- EVS and EAV (dotted circle) correspond to the chosen quality criteria (EVS) that are used in the verification analysis (EAV) resulted from the application of techniques defined in *(1)*;
- ASV, AEV, and APV (dotted circle) correspond to the preparation for applying validation (ASV and AEV) and then the execution (APV). Here, we could say that they are all together because for validation, only one category of technique is used (not like Verification, where *(2)* and EPV are two different categories of technique). Moreover, AVS is the definition of quality criteria to be used in the validation (APV).

We note that EDR - verification peer review (technique from *(1)*) is not present, being justified by the fact that the majority is *A* modality. For the variables ESV, EEV, EPV, ERR, ECR, and AEV that have the 6 highest contributions, we note that the first 5 variables (ESV, EEV, EPV, ERR, ECR) are related to verification. The last one (AEV) is only preparation for validation but had many *a* modality response because the HCI researchers suggest other examples of techniques.

Analyzing Figure 10 we confirm that there is a correspondence between HCI researchers (see clusters) and the items (see circles in Figure 9). The cluster of 5 (right side - Figure 10) is responsible for the most relevant result presented in Figure 9; because the cluster contains four *advanced HCI researchers* and only one *HCI researcher*, who gave a lot of suggestions (that means *a* modality) based on his industry experience. For the cluster of 9 (left side - Figure 10), the majority (6/9) are *HCI researcher*.

This agreement among some HCI researchers concerns their answers (in this case *partially agree*) for these variables. Contrariwise, the respondents did not all agree on the same justification, but in general, we obtained additional comments. For instance, regarding ESV variable we observe that *HCI researchers 8* and *18* are similar justification proposing verification methods as follows:

- **HCI researcher 8** – "Include verification by HCI researchers. I do not agree with validation by HCI researchers (selection of the method)."
- **HCI researcher 18 –** "Add other types of test. Use classical verification tests."

As a complementary justification, *HCI researcher 5* and *19* proposed to include other types of test:

- **HCI researcher 5 -** "Utility issues and tests should also be considered."
- **HCI researcher 19** – "Include test for accessibility. Consider task model."

Analyzing the justifications for ASV variable (in this case *partially agree*), we can note that the justifications of *HCI researchers* 8 and *16* proposed the inclusion of tests with the end user:

- **HCI researcher 8** – "I do not agree with validation by HCI researchers. Change the name of the category. Include validation with the end user (selection of the method)."
- **HCI researcher 16** – "I do not agree with validation by HCI researchers. Include tests with the end user."

HCI researchers 15 and *18* in their turn converged proposing the inclusion of communicability evaluation:

- **HCI researcher 15** – "Include communicability evaluation, user experience evaluation."
- **HCI researcher 18** – "Add other types of tests. Include communicability evaluation."

However, still for ASV other HCI researchers proposed other approaches:

- **HCI researcher 7** – "Include UI Evaluation methods (user review)."
- **HCI researcher 19** – "Include test for accessibility and validation by stakeholders."

The variables ESV, EEV, EPV, and EAV, are from the verification process that implied the exclusion of the proposed category and inclusion of a new category. This modification was suggested by two *advanced HCI researchers* (*HCI researchers 6 and 16*) who disagree with the proposals according to the following justifications:

- **HCI researcher 6** – "Considers classical verification tests of software engineering.
- **HCI researcher 16** – "I do not agree. Include verification by HCI researchers and classical tests of software engineering for the interactive part without the user. Use tasks model."

Analyzing Figure 11 and initial data, we can conclude that: (i) for the results from EPV variable related to *Industry* characteristic, 5/6 HCI researchers have industry experience, 4 were *advanced HCI researcher* with extensive experience in evaluation (verification and validation) systems, and they have *Other* as origin; (ii) ASV and AEV variables are related to the same specific goal (SG 1 of VAL – see Table 2); from ASV variable related to *Industry* characteristic, 5/6 HCI researchers have industry experience, 4 were *advanced HCI researcher* with extensive experience in evaluation (verification and validation) systems and they have *Other* as origin; and from AEV variable related to *Industry* characteristic, 7/9 HCI researchers have industry experience, 7 were *advanced HCI researcher* with experience in evaluation (verification and validation) systems and 4 of them have *Other* as origin.

5.5. Chernoff's faces and Summary

Authors acknowledge that MCA outputs (figures, tables that aid interpretation and texts describing the highlights of these figures and tables) are rather complex (but close to the initial information since all the data pieces are maintained in MCA input). Consequently, we try to show the results using a more understandable presentation. A principled psychological approach to data visualization is often warranted (Bertin, 1977) (Tufte, 1983) (Ware, 2000) (Lee et al., 2003) (Khan et al., 2015). Due to the presence of Agree/Disagree scale, we suggest using a rather well-known tool in statistics, i.e. Chernoff's faces (Chernoff, 1973).

This representation will be used because:

- after MCA (only the most discriminating variables and oppositions lined out by the main axes will be displayed). In most cases, this representation is employed to show the measurement variables with the initial statistical observations (Chernoff, 1973) and

(Johnson & Wichern, 1992); thus instead of having a single variable shown through one face part (mouth, nose, …), a part displays several variables with an identical value for a given variables subset, the subset emanating from MCA;

- with qualitative scales (with *A*, *a*, *D*, and *N* modalities). Although there are some examples with qualitative scales (Wainer & Thissen, 1981) and (Lee et al., 2003), this tool is mainly employed with quantitative variables (Chernoff, 1973) and (Johnson & Wichern, 1992);
- without the symmetric point of view (for instance, the two eyes show two different variables subsets). Here again, although there are some suggestions to use half of faces (Flury & Riedwyl, 1981) and (Tufte, 1983), the principle of the bilateral symmetry is more often present (Chernoff, 1973) and (Johnson & Wichern, 1992).

Combining these three aspects for Axes 1 and 2, Chernoff's faces are built in Figure 12. By analyzing Chernoff's faces, we can explain why 7 out of 9 variables related to vErification (Table 3) are the most discriminating.

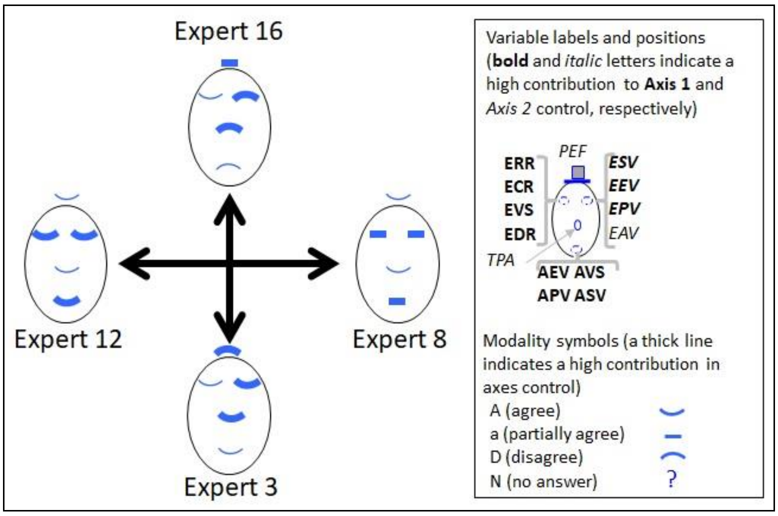

Figure 12. Chernoff's faces used to show the main result from MCA (see Figure 9 and Figure 10)

The presence of 4 out of 6 variables linked to vAlidation (AEV, AVS, APV, and ASV) with the same modality, concerns the use of a specific type of prototype. Combining these variables and modalities yields two opposed HCI researcher subsets (see Table 3) among which we can consider *HCI researchers 12* and *8* (Figure 12). The contribution of *HCI researcher 12* was in the practices related to requirements development. On the other hand, the main HCI domain of *HCI researcher 8* is evaluation, that explains his/her opinion as *partially agree* for the variables linked to vErification and vAlidation. We could have expected an opposition between modalities *A* and *D*, but this is not the case. The modality *D* appears for axis 2, essentially for the *HCI researcher 16*, who is a recognized HCI researcher in the evaluation of critical interactive systems.

As a summary, we see that the agreement about HCI approaches to support software engineering activities are mainly related to the Verification and Validation process areas of the CMMI-DEV. Moreover, the HCI researchers agree with the use of rapid prototypes for requirements elicitation. In addition, they also agree with the use of evolutionary prototypes (evolutionary versions of the system) for implementation (part of the Technical Solution process area), integration of products (part of Product Integration process area), verification

(Verification process area), and validation (Validation process area) of the system. We cannot conclude anything about the other process areas.

We note that although the gap between HCI and software engineering has been discussed more than fifteen years (Kazman et al., 2003) (John et al., 2004) (Folmer et al., 2006), there still a lack of agreement amongst HCI researchers. Analyzing engineering process areas regarding agreement among HCI researchers and their related HCI categories, we observe that:

- the category of standards and guidelines considers software engineering and HCI standards;
- functional prototypes have always been suggested by agile community as an important instrument to elicit and validate requirements;
- evaluation methods for HCI test include several approaches largely used in software engineering;
- usability evaluation is one of the approaches most used in HCI practice;
- evaluation methods for review (e.g. peer review) consider some approaches that were derived from software engineering that maybe facilitates integration of review with the software engineering processes.

We conclude that for HCI approaches for the verification and validation processes, there is an agreement among HCI researchers. For several HCI approaches for requirements development, technical solution, and product integration processes, there is some agreement among HCI researchers. However, we find a lack of agreement among HCI researchers for many other HCI approaches, in the requirements development and technical solution processes.

These findings are consistent with the HCI literature (Ogunyemi, Lamas, Lárusdóttir, & Loizides, 2018) which identified that some studies in this domain are focused on evaluation research; and HCI techniques are used in requirements development and evaluation (validation) phases.

We explain the lack of agreement among HCI researchers for several HCI approaches as follows:

- the HCI domain is evolving very fast with many methods, techniques, models, patterns, and so on; see for instance, the content of HCI journals and the last conference proceedings of Conference on Human Factors in Computing System (CHI), International Conference on Human-Computer Interaction (INTERACT), International Conference on Human-Computer Interaction (HCI International), International Conference on Tangible, Embedded and Embodied Interaction (TEI), International Conference on Human-Computer Interaction with Mobile Devices and Services (MOBILEHCI) or Designing Interactive Systems (DIS);
- new application domains (Internet of Things (IoT), tangible interaction, big data, games/serious games, and so on) lead to new analysis and design issues; iii) new technical solutions progressively appear or evolve (augmented and virtual reality, advanced networks, many new types of interactive devices, and so on;
- we believe that HCI and software engineering are not considered very close to most HCI researchers (including maybe some respondents of our survey).

As a result, it is not easy for HCI researchers to have an accurate and stabilized view of the entire domain.

6. Threats to Validity

We analyzed four threats of validity proposed by (Wohlin et al., 2012): construct validity, internal validity, conclusion validity, and external validity.

Construct validity "is concerned with the relation between theory and observation" (Wohlin et al., 2012). This threat concerns the construction and the use of the questionnaire in this study. To minimize this threat, we constructed the questionnaire using the original text extracted (specific goals and specific practices) from the official documentation of CMMI-DEV. The HCI categories were collected from literature and pre-validated by one of the authors of this study. Moreover, two authors who conducted the interviews have good knowledge about CMMI-DEV; and we had the official documentation of CMMI-DEV over the interviews. Therefore, we consider this risk under control.

Threats to **internal validity** "are influences that can affect the independent variable concerning to causality, without the researchers' knowledge" (Wohlin et al., 2012). In this study, the threat is associated with the participants (HCI researchers). The first group of participants was selected by convenience from the professional network of one of the authors. After that, these participants suggested other names following a pre-defined profile (as presented in section 4.1). The literature shows that some studies have used small samples to obtain researcher feedback. For instance:

- (Dyba, 2000) had 11 researchers who revised a process;
- (Huart et al., 2004) present a study about the evaluation of several multimedia applications made by four researchers using Cognitive Walkthrough;
- (Beecham et al., 2005) had 20 researchers who evaluate their requirements process improvement model;
- (Følstad et al., 2010) present a study about the usability inspection performance with the participation of fifteen work-domain researchers and twelve usability researchers;
- (Gil Urrutia et al., 2017) present a study where they search how UX professionals (seventeen researchers) organize their cognitions around some criteria.

Therefore, we accepted this risk because it is normal to have small size when the participants are domain researchers.

Another internal validity threat is the HCI researchers' knowledge about the HCI categories, their examples, and about HCI good practices (international standards, professional certification for usability professionals, etc.). We assumed that the participants knew the proposed HCI categories and they know how to apply the approaches in the development of interactive systems. To mitigate this threat, we selected only participants that have experience (academic and/or industrial) in the HCI domain and have a Ph.D. degree which means knowing the approaches and where those approaches are used in interactive system development even if they did not work in practice. In addition, we decided that it was not necessary to be familiar with CMMI-DEV since the practices of engineering process areas are described in the questionnaire. However, we could not assure that all participants have complete knowledge of the use of approaches in the development of interactive systems. Since we were working with researchers with large experience, we decide to accept this risk.

The **conclusion validity** "is concerned with the relationship between the treatment and the outcome" (Wohlin et al., 2012). The conclusion threat is the relation between the HCI categories, and each specific practice evaluated by participants. To reduce this risk, we decided to perform interviews individually and not to use a survey. In this way, we could clarify each question by the participants about the objective of the study, the CMMI-DEV, and the HCI categories. In addition, when the participants partially agreed or disagreed with one or more propositions, they were asked to justify their opinion and include any other proposals (when necessary). Moreover, we considered two groups of participants (from France, majority, and other countries).

Threats to **external validity** are "concerned with generalization" (Wohlin et al., 2012). The result could be biased if participants come only from one domain (HCI). To mitigate this risk, we interviewed participants with different expertise in HCI and have experience recognized by HCI community (e.g. program chair or member of the program committee of HCI conferences, editor of journals and members of HCI associations). In addition, we have invited HCI researchers that are well known for working on different technologies (e.g. web applications, information systems, critical systems, tabletop applications, and so on). However, we are aware that even with all these mitigation procedures; it is not possible to generalize our result for all communities, considering the size of the sample and representativeness of countries.

7. Conclusion

Taking advantage of the data collected for a previous study performed about the identification of HCI approaches to support SPCM models, this paper presents the degree of agreement among HCI researchers about software engineering activities.

For this study we considered the data of the interview with 20 HCI researchers from the HCI domain. For five months, we performed a multivariate analysis of the results of these interviews that is presented in this paper.

From the analysis of this study, we can conclude that although we had several proposals for HCI approaches to support software engineering practices, the proposals for the *Verification* and *Validation* process areas have the best agreement among the HCI researchers. These process areas are much closer because several approaches were used by both Software Engineering (SE) and Human-Computer Interaction (HCI) domains; some researchers who contributed to the results worked mainly in HCI evaluation.

The qualitative analysis indicates that the use of prototypes for the last phases of the software development is justified only if they are considered as evolutionary system versions. We identified a lack of agreement among HCI researchers for some HCI approaches in the requirements development, technical solution, and product integration process areas; this is due to significant individual differences.

The result obtained from the statistics analysis confirms that SE and HCI are not yet very close for all processes of the software development. We conclude that although a lot of efforts have been made in industry to use HCI practices integrated in the development of software systems, the academic and industry environments need to enlarge still more their investment in the integration of HCI and SE practices. Regarding the academic environment, we need to propose training that integrates both domains (SE and HCI) and their skills.

For the industrial environment, we already have international standards that propose human-centered design concepts (e.g. ISO 9241-210:2010 (ISO, 2010) and ISO 9241-220:2019 (ISO, 2019)) to support the system development. It is also important to consider HCI good practices found in international standards (such as ISO 9241-100 series of standards, and ISO/TR 16982:2002 (ISO, 2002)).

In addition, the enterprises that develop systems based on Agile approaches could explore the use of guidelines (for instance, to Scrum and CMMI (CMMI Product Team, 2016)).

Finally, in order to better generalize the results proposed in this study, new interviews with other representative HCI researchers from other countries are envisaged as future research.

Acknowledgements

The authors would like to acknowledge the financial support granted by CAPES – Science without Borders Program (DOC-PLENO 99999.013381/2013-00). They also thank the HCI researchers who participated in this study, and the anonymous reviewers for their numerous constructive remarks.

References

Beecham, S., Hall, T., Britton, C., Cottee, M., & Rainer, A. (2005). Using an expert panel to validate a requirements process improvement model. *Journal of Systems and Software*, *76*(3), 251–275. http://dx.doi.org/10.1016/j.jss.2004.06.004

Beh, E. J., & Lombardo, R. (2014). *Correspondence analysis: Theory, practice and new strategies*. John Wiley & Sons Inc.

Benzécri, J.-P. (1992). *Correspondence analysis handbook*. CRC Press.

Bertin, J. (1977). *La graphique et le traitement graphique de l'information*. Flammarion.

Bevan, N. (2009). International Standards for Usability Should Be More Widely Used. *Journal Usability Studies*, *4*(3), 106–113.

Chernoff, H. (1973). The Use of Faces to Represent Points in K-Dimensional Space Graphically. *Journal of the American Statistical Association*, *68*(342), 361–368.

CMMI Product Team. (2010). *CMMI® for Development* (CMU/SEI-2010-TR-033; Version 1.3). Software Engineering Institute, Carnegie Mellon University. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=9661

CMMI Product Team. (2016). *A Guide to Scrum and CMMI®: Improving Agile Performance with CMMI* (p. 132). CMMI Institute. https://cmmiinstitute.com/resource-files/public/marketing/document/a-guide-to-scrum-and-cmmi®-improving-agile-perfor

Cornillon, P. A. (2018). *R pour la statistique et la science des données*. Presses Universitaires de Rennes.

Crawley, M. J. (2007). *The R book*. Wiley.

Dyba, T. (2000). An Instrument for Measuring the Key Factors of Success in Software Process Improvement. *Empirical Software Engineering*, *5*(4), 357–390. https://doi.org/10.1023/A:1009800404137

Flury, B., & Riedwyl, H. (1981). Graphical Representation of Multivariate Data by Means of Asymmetrical Faces. *Journal of the American Statistical Association*, *76*(376), 757–765.

Folmer, E., Welie, M. van, & Bosch, J. (2006). Bridging patterns: An approach to bridge gaps between SE and HCI. *Information and Software Technology*, *48*(2), 69–89. https://doi.org/10.1016/j.infsof.2005.02.005

Følstad, A., Anda, B. C. D., & Sjøberg, D. I. K. (2010). The usability inspection performance of work-domain experts: An empirical study. *Interacting with Computers*, *22*(2), 75–87. https://doi.org/10.1016/j.intcom.2009.09.001

Friendly, M., & Meyer, D. (2016). *Discrete Data Analysis With R: Visualization and Modeling Techniques for Categorical and Count Data* (Har/Psc). Productivity Press.

Gil Urrutia, J. I., Brangier, E., & Cessat, L. (2017). Is a Holistic Criteria-Based Approach Possible in User Experience? In A. Marcus & W. Wang (Eds.), *Design, User Experience, and Usability: Theory, Methodology, and Management* (Vol. 10288, pp. 395–409). Springer International Publishing. https://doi.org/10.1007/978-3-319-58634-2_29

Gonçalves, T. G., Oliveira, K. M., & Kolski, C. (2018). Identifying HCI Approaches to support CMMI-DEV for Interactive System Development. *Computer Standards & Interfaces*, *58*, 53–86. https://doi.org/10.1016/j.csi.2017.12.003

Gonçalves, T. G., Oliveira, K. M., & Kolski, C. (2016). HCI engineering integrated with capability maturity models: A study focused on requirements development. *Tenth IEEE International Conference on Research Challenges in Information Science - RCIS*, 633–644. https://doi.org/10.1109/RCIS.2016.7549319

Ham, D.-H. (2014). A model-based framework for classifying and diagnosing usability problems. *Cognition, Technology & Work*, *16*(3), 373–388. https://doi.org/10.1007/s10111-013-0267-6

Hao, H. M., & Jaafar, A. (2011). Usability in Practice: Perception and Practicality of Management and Practitioners. *Proceedings of International Conference on Pattern Analysis and Intelligence Robotics*, *2*, 211–216. https://doi.org/10.1109/ICPAIR.2011.5976928

Huart, J., Kolski, C., & Sagar, M. (2004). Evaluation of multimedia applications using inspection methods: The Cognitive Walkthrough case. *Interacting with Computers*, *16*(2), 183–215. https://doi.org/10.1016/j.intcom.2003.12.005

ISO. (2002). *Ergonomics of human-system interaction – Usability methods supporting human centred design (ISO/TR 16982:2002)*.

ISO. (2008). *Ergonomics of human-system interaction—Part 151: Guidance on World Wide Web user interfaces (ISO 9241-151:2008)*.

ISO. (2010). *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (ISO 9241-210:2010)*.

ISO. (2016). *Ergonomics of human-system interaction—Part 161: Guidance on visual user-interface elements (ISO 9241-161:2016)*.

ISO. (2017a). *Ergonomics of human-system interaction—Part 112: Principles for the presentation of information (ISO 9241-112:2017)*.

ISO. (2017b). *Ergonomics of human-system interaction—Part 125: Guidance on visual presentation of information (ISO 9241-125:2017)*.

ISO. (2018). *Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts (ISO 9241-11:2018)*.

ISO. (2019). *Ergonomics of human-system interaction – Part 220: Processes for enabling, executing and assessing human-centred design within organizations (ISO 9241-220:2019)*.

John, B. E., Bass, L., Kazman, R., & Chen, E. (2004). Identifying Gaps Between HCI, Software Engineering, and Design, and Boundary Objects to Bridge Them. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1723–1724. https://doi.org/10.1145/985921.986201

Johnson, R. A., & Wichern, D. W. (1992). *Applied multivariate statistical analysis*. Prentice-Hall.

Jokela, T., Siponen, M., Hirasawa, N., & Earthy, J. (2006). A survey of usability capability maturity models: Implications for practice and research. *Behaviour & Information Technology*, *25*(3), 263 – 282. https://doi.org/10.1080/01449290500168079

Kazman, R., Bass, L., & Bosch, J. (2003). Bridging the Gaps Between Software Engineering and Human-Computer Interaction. *Proceedings of the 25th International Conference on Software Engineering*, 777–778.

Khan, A., Breslav, S., Glueck, M., & Hornbæk, K. (2015). Benefits of visualization in the Mammography Problem. *International Journal of Human-Computer Studies*, *83*, 94–113. http://dx.doi.org/10.1016/j.ijhcs.2015.07.001

Lee, M. D., Butavicius, M. A., & Reilly, R. E. (2003). Visualizations of binary data: A comparative evaluation. *International Journal of Human-Computer Studies*, *59*(5), 569–602. http://dx.doi.org/10.1016/S1071-5819(03)00082-X

Mohamed, M. A., Chakraborty, J., & Dehlinger, J. (2017). Trading off usability and security in user interface design through mental models. *Behaviour & Information Technology*, *36*(5), 493–516. https://doi.org/10.1080/0144929X.2016.1262897

Nishisato, S. (2007). *Multidimentional nonlinear descriptive analysis*. Chapman & Hall/CRC.

Ogunyemi, A. A., Lamas, D., Adagunodo, E. R., Loizides, F., & Rosa, I. B. D. (2016). Theory, Practice and Policy: An Inquiry into the Uptake of HCI Practices in the Software Industry of a Developing Country. *International Journal of Human–Computer Interaction*, *32*(9), 665–681. https://doi.org/10.1080/10447318.2016.1186306

Ogunyemi, A. A., Lamas, D., Lárusdóttir, M. K., & Loizides, F. (2018). A Systematic Mapping Study of HCI Practice Research. *International Journal of Human–Computer Interaction*, 1–26. https://doi.org/10.1080/10447318.2018.1541544

Ogunyemi, A., Lamas, D., & Eze, E. (2018). Exploring the State of Human-centred Design Practice in Software Development Companies: A Cross-Case Analysis of Three Nigerian Software Companies. *Interacting with Computers*, *30*(5), 444–467. https://doi.org/10.1093/iwc/iwy020

Scheiber, F., Wruk, D., Oberg, A., Britsch, J., Woywode, M., Maedche, A., Kahrau, F., Meth, H., Wallach, D., & Plach, M. (2012). Software Usability in Small and Medium Sized Enterprises in Germany: An Empirical

Study. In A. Maedche, A. Botzenhardt, & L. Neer (Eds.), *Software for People: Fundamentals, Trends and Best Practices* (pp. 39–52). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31371-4_3

Schutte, P. C. (2017). How to make the most of your human: Design considerations for human–machine interactions. *Cognition, Technology & Work, 19*(2), 233–249. https://doi.org/10.1007/s10111-017-0418-2

Sears, A., & Jacko, J. A. (2009). *Human-Computer Interaction: Design issues, solutions, and applications*. CRC Press.

Seffah, A., Gulliksen, J., & Desmarais, M. C. (2005). *Human-Centered Software Engineering—Integrating Usability in the Software Development Lifecycle* (Vol. 8). Springer Netherlands.

Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures* (4th ed.). Chapman & Hall/CRC.

Shin, Y., Im, C., Oh, H., & Kim, J. (2017). Design for experience innovation: Understanding user experience in new product development. *Behaviour & Information Technology, 36*(12), 1218–1234. https://doi.org/10.1080/0144929X.2017.1368709

Tufte, E. (1983). *The Visual Display of Quantitative Information*. Graphics Press.

Wainer, H., & Thissen, D. (1981). Graphical data analysis. *Annual Review of Psychology, 32*, 191–241.

Ware, C. (2000). *Information Visualization: Perception for Design*. Morgan Kauffman.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer-Berlin Heidelberg.