A Tool for Getting Cultural Differences in HCI

Rüdiger Heimgärtner
IUIC (Intercultural User Interface Consulting)
Germany

1. Introduction

The "Intercultural Interaction Analysis" tool (IIA tool) was developed to obtain data regarding cultural differences in HCI. The main objective of the IIA tool is to observe and analyze the interaction behavior of users from different cultures with a computer system to determine different interaction patterns according to their cultural background. Culture influences the interaction of the user with the computer because of the movement of the user in a cultural surrounding (Röse, 2002). To locate and find out the kind of different interaction behavior of the users from different cultural groups (at national level (country) between Chinese and German user first because of the high cultural distance) the interaction behavior of the users with the computer will be observed and detected. The objective is to be able to draw inferences regarding differences of the cultural imprint of users by analyzing the interaction behavior of those users with a computer system to get knowledge that is relevant for intercultural user interface design and a necessary precondition for cultural adaptive systems (Heimgärtner, 2006). E.g. the right number and arrangement of information units is very important for an application whose display is very small and at the same time the mental workload of the user has to be as low as possible (e.g. driver navigation systems).

2. Designing a Tool for the Analysis of Cultural Differences in HCI (IIA Tool)

Research of literature showed that there are no adequate methods for determining cross-cultural differences in interaction aspects of human machine interaction (HMI) and none for driver navigation systems. For doing this task on PC’s, there are some tools like:

- UserZoom (Recording, analyzing and visualization of online studies)
- ObSys (Recording and visualization of windows messages) (cf. Gellner & Forbrig, 2003)
- INTERACT (Coding and visualization of user behavior) (cf. Mangold, 2005)
- REVISER (Automatic Criteria Oriented Usability Evaluation of Interactive Systems, cf. Hamacher, 2006)
- Noldus, SnagIt, Morae, A-Prompt, Leo, etc.

All the existing tools provide some functionality for (remote) usability tests and interaction behavior measurement. Nevertheless, I had to develop my own tool for this purpose, because this task presupposes intercultural usability metrics, i.e. a Cross-Cultural Usability Metric Trace Model (CCUMTM) or even better a Cultural HMI Metric Model (CHMIMM), which none of the existing tools offer explicitly for this purpose (because the parameters and
CCUMTM did not exist). This needs knowledge about variables depending on culture, which could not have been implemented in the existing tools. On the one hand, they are not known up to now. On the other hand, the architecture of the existing tools cannot be changed such that the potential cultural parameters can be determined by tests with the tools.

My theoretical reflections and deductions from literature led to a hypothetical model of intercultural variables (IV model) for the HMI design. It must be distinguished between variables, that can be determined at runtime, and variables, whose values must be determined in design phase to provide them for the runtime system. A benchmark test of systems from different countries with similar functions can help to determine differences in HMI. Furthermore, the interaction of cultural different users doing the same task should be observed (using the same test conditions i.e. the same hard and software, environment conditions, language, experience of using the system as well as the same test tasks). Helpful are also data of diagnose, debugging and HMI event triggering during usage of the system summarized in the Usability Metric Trace Model (UMTM). These data can be logged during usability tests according to certain user tasks. The evaluation of the collected data using statistical methods should show, which of the potential variables depend on culture (potential cultural interaction indicators (PCII’s)). Having this knowledge, the UMTM can be optimized and verified empirically by further experiments within usability tests to get the cross-cultural UMTM (CCUMTM). This requires several development loops within integrative design.

To motivate the user to interact with the computer and to verify the postulated hypotheses, adequate task scenarios have been developed and implemented into the IIA tool. Even if the architecture of this new tool follows in some respect the already existing tools, it has been developed from the scratch because the existing tools did not measure intercultural interaction behavior according to driver navigation use cases which was a main requirement getting budget for developing the IIA tool. The resulting tool provides data collection, analysis, and evaluation for intercultural interaction analysis in HCI:

- recording, analysis and visualization of user interaction behavior and preferences
- localized tool for intercultural usability testing using use cases that are comprehensive in different cultures
- integration of usability evaluation techniques and all interaction levels according to the acting level model (cf. Herczeg, 2005).
- qualitative judgments by quantitative results (optimization of test validity and test reliability).

The preparation of the collected data takes place mostly automatically by the IIA data collection tool, which saves much time, costs, and effort. The collected data is partly quantitative (related to all test persons, e.g. like the mean of a Likert scale) and partly qualitative (related to one single test person, e.g. answering open questions) (cf. De la Cruz et al., 2005). Moreover, the collected data sets have standard format so that anyone can perform own statistical analyses. This also means that the results of studies using the IIA tool are verifiable because they can be reproduced using the IIA tool. The data will be stored in databases in formats (CSV, MDB) that are immediately usable by the IIA analysis tool, and, which also conducts possible subsequent converting and data preparation. Hence, statistic programs like SPSS, AMOS, and neural network can be deployed to do descriptive or explanatory statistics, correlations, and explorative or confirmatory factor analysis, to
explore cultural differences in the user interaction as well as to find a cultural interaction model using structural equal models. The data evaluation module enables classification with neural networks to cross-validate the results from data analysis. In future, it will be extended such, that it is possible to evaluate the analysis on the fly during data collection. The quantitative studies should reveal trends for the investigated cultures regarding the interaction behavior with the computer. Data mining methods and statistics e.g. cluster analysis for classification or linear regression for correlations can be exploited to find correlations between recorded cross-cultural user interaction values and values of the cultural variables (cf. Kamentz & Mandl, 2003).

Delphi was used to create a software tool, which can be installed online using the Internet as well as offline via CD. To avoid downloading and interaction delays, the IIA tool has been implemented also in one single executable program file on a server to be downloaded onto the local hard disk of the users worldwide because the tool has to measure the interaction behavior of the user during the online tests correctly and comparably. A huge amount of valid data can be collected rapidly and easily worldwide online via internet or intranet. Besides, the Delphi IDE allows transforming new HMI concepts and test cases very quickly into good-looking prototypes that can be tested very soon in the development process. E.g., some hypotheses could have been confirmed quantitatively addressing many test users online using the IIA tool within one month (implementing the use cases as well as doing data collection and data analysis). Hence, using the IIA tool means rapid use case design, i.e. real-time prototyping of user interfaces for different cultures.

3. Implementation of Test Tasks and the UMTM

The IIA tool has been developed to be able to determine the intercultural differences in the basic principles of HMI as well as in the use cases related to special products (e.g. driver navigation systems). Hence, the results can be general guidelines for every intercultural HMI development as well as context specific recommendations for the design of special products. The intercultural interaction analysis tool provides an implementation of the UMTM and therefore the ability to determine the peculiarities and values of the specified intercultural variables. Thereby, the IIA tool serves to analyze cultural differences in HMI.

The following information scientific parameters (information related dimensions) can be determined quantitatively:

- Information density (spatial distance between informational units)
- Information speed (time distance between informational units to be presented)
- Information frequency (number of presented informational units per time unit)
- Interaction frequency (number of initialized interaction steps per time unit)
- Interaction speed (time distance between interaction steps)

Not all PCII’s from IV model and UMTM could have been implemented into the IIA data collection module because of time and budget restrictions. Only the most promising PCII’s requiring the least integrating effort to the test system have been implemented. Nevertheless, more than one hundred potentially culturally sensitive variables in HMI have been implemented into the IIA tool, and applied by measuring the interaction behavior of the test persons with a personal computer system in relation to the culture (as presented in table 1).
| Measured variables in the single test tasks | URD (user requirement design) test task | MD (map display) test task | MG (maneuver guidance) test task | IO (information order) test task | INE (interaction exactness) test task | INS (interaction speed) test task | QUES (questionnaire) test task | IH (information hierarchy) test task | UV (uncertainty avoidance) test task |
|--------------------------------------------|----------------------------------------|---------------------------|---------------------------------|---------------------------------|----------------------------------|-------------------------------|---------------------------|---------------------------------|----------------------------------|
| | PositionXBegin(URD), PositionYBegin(URD), PositionXEnd(URD), PositionXBack(URD), PositionYBack(URD), PositionXNext(URD), PositionYNext(URD), PositionXEnd(URD), PositionYEnd(URD), PositionXReady(URD), PositionYReady(URD), PositionXDisplay(URD), PositionYDisplay(URD), PositionXListbox(URD), PositionYListbox(URD), PositionXStatus(URD), PositionYStatus(URD) | NumberOfTextures(MD), NumberOfPOI(MD), NumberOfStreetNames(MD), NumberOfStreets(MD), NumberOfManoeuver(MD), NumberOfRestaurants(MD) | MessageDistance(MG), DisplayDuration(MG), CarSpeed(MG) | InformationorderNumber(IO), InformationorderOrder(IO), FactorOfUnorder(IO), PixelOfUnorder(IO), PixelOverlapping(IO), CoverageFactor(IO), PixelSize(IO), DistanceImageMargin(IO), PixelDistance(IO) | InteractionexactnessSpeed(INE), InteractionexactnessExactness(INE) | InteractionspeedExactness(INS), InteractionspeedSpeed(INS) | ChangeValueEndQues(QUES) | InformationhierarchyNumber(IH) | UncertaintyAvoidanceValue(UV) |

| Measured variables at each test task | TestTaskDuration, TotalDialogTime, NumberOfErrorClicks, NumberOfMouseClicks, EnteredChars(where possible) |
|------------------------------------|--------------------------------------------------|

| Measured variables over the whole test session | TestDuration, TotalDialogTime, MaximalOpenTasks, NumberOfScrolls, AllMouseClicks, NumberOfErrorClicks, NumberOfMouseClicks, MouseLeftUps, MouseLeftDowns, ClickDistance, ClickDuration, NumberOfMouseMoves, MouseMoveDistance, NumberOfAgentMoves, NumberOfAgentHides, NumberOfShowMessages, NumberOfONOs, NumberOfYESs, NumberOfAcknowledgedMessages, NumberOfRefusedMessages, Lex (syntactical entries), Sem (semantical entries), (Interaction-)Breaks0ms, Breaks1ms, Breaks10ms, Breaks100ms, Breaks1s, Breaks10s, Breaks100s, Breaks1000s, Breaks10000s |

| Measured variables before the test session | OpenTasksBeforeTest |

Table 1. Implemented variables from the UMTM in the IIA tool (the test tasks will be explained below in detail)

As mentioned above, the IIA tool allows the measurement of numerical values like information speed, information density, and interaction speed in relation to the user. These are hypothetically correlated to cultural variables concerning the surface like number or position of pictures in the layout or affecting interaction like frequency of voice guidance. Every one of the test tasks serves to investigate other cultural aspects of HCI. The test setting within the IIA tool contains two scenarios:
• an abstract scenario with tasks for general usage of widgets and
• a concrete scenario with tasks for using a driver navigation system.

In the first scenario, the user uses certain widgets. Those tasks can only be done by persons that have seen and used a PC before. The second scenario takes into account concrete use cases from driver navigation systems. The requirements of those tests are that the user has some knowledge and interaction experience about driver navigation systems as well as about PCs. The results of the abstract test cases are expected to be valid for HMI design in general because the context of usage is eliminated by abstract test settings, which are independent from actual use cases. The simulation of special use cases within the IIA tool can show usability problems and differences in user interaction behavior (similar to “paper mock-ups”). The test tasks are localized at technical and linguistically level, but are semantically identical for all users, so that participants of many different cultures can do the IIA test. Hence, the study can be extended from Chinese and German to other cultures in different countries by using the same (localized) test tool. Both abstract and special test cases have been implemented in this way as test scenarios into the IIA data collection tool in order to obtain results for the intercultural HMI design (cf. Heimgärtnert, 2005). To transfer the results of general test cases in the abstract test settings to driver navigation systems, special use cases had been implemented as test scenarios in the test tool. E.g., the hypothesis “there is a high correlation of high information density to relationship-oriented cultures such as China” should be confirmable by adjusting more points of interest (POI) by Chinese users compared to German users. So, the use case “map display” was simulated by the map display test task to measure the number of pieces of information on the map display regarding information density (e.g. restaurants, streets, POI, etc.) (cf. figure 1).

Fig. 1. Screenshot of the “map display test task” during the test session with the IIA data collection module. The user can define the amount of information in the map display by adjusting the scroll bars. The test tool records the values of the slide bars set up by the user.
Based on this principle, the test tool can also be used to investigate the values of other cultural variables like widget positions, menu structure, layout structure, interaction speed, speed of information input, dialog structure, etc. The test with the IIA tool was designed to help to reveal the empirical truth to such questions. Some of these aspects and use cases will be explained in more detail in this section to get an impression of the possible relationship between the usage of the system by the user and their cultural background. Along the implemented use case “map display” in the map display test task shown in figure 1, another important use case of driver navigation systems “maneuver guidance” has implemented as maneuver guidance test task into the IIA data collection module. The test user has to adjust the number and the time distance of the maneuver advice messages on the screen concerning frequency and speed of information (cf. figure 2).

Fig. 2. Maneuver Guidance Test Task. The test person can use the sliders to select the car speed (indicated by the red rectangle), the duration of displaying the maneuver advice as well as the time distance of the given hints.

Another variable is e.g. measuring the acceptance of the “life-like” character "Merlin". According to Prendinger & Ishizuka 2004, such avatars can reduce stress during interaction with the user. Hence, the agent “Merlin” was implemented in the IIA tool to offer his help every 30 seconds (cf. figure 3). On the one hand, according to cultural dimensions (cf. Marcus & Baumgartner, 2005), which describe the behavior of human beings of different

---

1 The virtual assistant „Merlin“ is part of the interactive help system of Microsoft Office™.
cultures, like high uncertainty avoidance or high task orientation, it was expected that German users switch off the avatar very soon (compared to Chinese users), because they do fear uncertain situations (cf. Hofstede et al., 2005). Furthermore, they do not like to be distracted from achieving the current task (cf. Halpin et al., 1957). On the other hand, if applying the cultural dimension of face saving, it should be the other way around. If Chinese users make use of help very often, they would lose their face (cf. Victor, 1997; Honold, 2000).

The interaction speed test task is very abstract and is not related to DNS. Figure 4 shows the graphical user interface (GUI) for this test task. The user has to click away 16 randomly arranged dots at the screen to be able to measure interaction speed and sequentiality (clicking order). Similar to this test task is the interaction exactness test task, which measures the same parameters, but displays the points sequentially (to measure the clicking exactness, i.e. a deviation factor from the middle of the dots). Thereby, the following PCII’s can be measured:

- Average time from clicking off one dot to another.
- Sequence of clicking off the dots.
- Exactness of clicking the dot in the middle.
- Number of interaction breaks during doing the task.
- Time period between information presentation and next user interaction with the system (user response time).
- Test task duration.
In an additional test task, the user has the possibility to specify his requirements for widget position directly visually by designing the layout of the GUI e.g. by changing the widget position within the user requirement design (URD) test task. Figure 5 shows the main part of the GUI of the URD test task. Here, the following PCII’s can be determined:

- Position of widgets.
- Duration of drag and drop process.
- Moving speed.
- Sequence of handling the widget.
- Number of function initiations (e.g. during testing the widget functions after finishing their arrangement).
- Sequence of function initiations (e.g. during testing the widget functions after finishing their arrangement).

During the whole test session, the IIA tool records the interaction between user and system, e.g. mouse moves, clicks, interaction breaks, or the values and changing’s of the slide bars set up by the users in order to analyze the interactional patterns of the users of different culture. Thereby, all levels of the interaction model (physical, lexical, syntactical, semantic, pragmatic, and intentional) necessary for dialog design can be analyzed (cf. Herczeg, 2005). Figure 6 shows a part of a course of interaction of a user with the system during the test.
session represented by some parameters like mouse moves or mouse clicks as well as keyboard presses (at y-axis) displayed over time (at x-axis).

Fig. 5. Main part of the GUI of the URD Test Task

Fig. 6. Part of a Course of Interaction (of a user with the system during a test session)
4. IIA Tool Setup, Test Setting and Usage

To motivate the user to interact with the computer and to test the hypotheses, test tasks have been developed and implemented into the IIA tool as described in the last section, which the user has to work on. Figure 7 shows the IIA test procedure containing the sequence of tasks presented to the test participant (the brackets embraces the file names of the source code of the modules written in Delphi7).

Fig. 7. Test procedure of the IIA test

A user test session with the IIA data collection module comprises five parts: collection of demographic data, test tasks, (cultural) value survey module (VSM94) questionnaire from Hofstede 1994, evaluation of results by the user, and debriefing questionnaire. The method to ask many users online by letting them do special test use cases and to collect the qualitative data (user preferences) emerged by this process quantitatively, has been used for Chinese (C), English (E) and German (G) speaking employees of SiemensVDO (SV) (now Continental) worldwide by an automated online data collection using the IIA data collection module to get cultural differences in HMI. After the start of the IIA data collection module, firstly, the user has to select his preferred test language (figure 8).
Fig. 8. The user can choose the language in which the test takes place (Chinese, German, or English)

Afterwards, greetings and a legend will be presented followed by a declaration of consent by the user that the collected data from the user may be used within the research project (figure 9). If the user disagrees, no personal data may be collected: the data collection will be anonymous.

Fig. 9. Test introduction and agreement to use personal data
The demographic “questionnaire” delivers the usual knowledge of demographic research especially about the cultural background of the user (like mother tongue, languages, nationality, residence in foreign countries, highest education, job description, age and PC experience) (figure 10).

However, in this case, the demographic “questionnaire” is already a special test task, recording also many parameters regarding the interaction behavior of the user with the system.

- Sequence of asking the questions.
- Number of dialog steps to finish the test task.
- Number of interactions during doing the test task e.g. number of using optional functions and help initiations, color settings, mouse moves or clicks and drop downs.
- Length of interaction breaks during doing the test task.
- Number of premature trials to go on to the next test task because the user meant he has finished the current task already. It is assumed that (C)>(G), because C has lower uncertainty avoidance than (G).
• Number of help usage. The user can press a help button to get a hint about to do the test task. It is assumed in literature (cf. Honold, 2000), that Chinese users do not use this button as often as German users because of fearing to lose their face.

• Number of initiating optional functions supposed to be high for (G) because of the wish to work very accurately.

• Straightness of mouse moving direction is assumed to be linear for (G) because of high task-orientedness.

• Speed of mouse movement: probably higher for (C) than for (G) because of low uncertainty avoidance and high communication speed.

• Jerkiness of mouse movements concerning affectivity and emotionality: (C) higher than (G) because of their relationship-orientedness.

• Number of language switching probably higher for (C) because of cultural interest and openness as well as multilingualism by relationship-orientedness and collectivistic attitude.

• Number of dialog steps assumed to be lower for (G) than for (C) because of task-orientedness.

• Test duration can be both: (G) > (C) by doing tasks very exactly because of task-orientedness but also the other way around: (C) > (G) by discussing the tasks with other people because of relationship-orientedness.

• For all number of key presses like usage of return-key-presses and usage of keyboard (number and kind of key presses), (C) > (G) is expected because of high interaction speed coming from low uncertainty avoidance as well as high communication speed and density by relationship-orientedness.

• Number of sounds, words, sentences, and utterances higher for (C) e.g. because of higher affectivity and relationship-orientedness.

• Duration of selection (e.g. combo box) is expected higher for (G) than for (C) because of degree of reflection (R) and interaction exactness (hit exactness at motoric selection).

• Time between “Mouse-Move-Over-Widget and Click onto Widget”: (G) > (C) because high degree of reflection and low interaction speed.

• Length of stay with mouse at widgets: (G) > (C) because of higher degree of reflection (R) as well as low interaction speed.

• Double click speed: higher for (C) than for (G) because of interaction speed, uncertainty avoidance (UV) and affectivity.

• Entering speed (e.g. on the keyboard): (C) > (G) because of interaction speed, (UV) and affectivity.

• Times between “selecting” and “using” (G)> (C) because of interaction speed, (UV) and (R).

• Sequence of user actions (e.g. „Selecting“): (G) > (C) because of action chain theory, i.e. high information sequentiality because of orientation to plan – avoiding coincidences – and doing things sequentially according to mono-chrone understanding of time as well as high uncertainty avoidance exposing a linear cognitive style.

• Number of backspace usage (number of wrong entering) and error clicks (= senseless or useless mouse clicks): (C) > (G) because of low (UV) and high interaction speed paired with high impatience or desire to get fast feedback (all initiations are expected to get immediate reactions).
Fig. 11. Question in the IIA data collection module

| Parameter          | Example                          |
|--------------------|----------------------------------|
| ID                 | 47                               |
| Number             | 42                               |
| Group              | 2                                |
| Category           | DNS                              |
| Position           | 4                                |
| Source             | Brown                            |
| Headline           | Driver navigation system         |
| Question           | How polite should be a DNS?      |
| Scale              | Interval                         |
| Scale size         | 0.100                            |
| Attributes         | polite honestly, polite-euphemistic |
| Layout             | Vertical                         |
| Result             | 54                               |
| Reason necessary   | False                            |
| Reason box headline| Please give a reason             |
| Reason             | no idea                          |
| Show values        | False                            |
| Priority           | 4                                |
| Without question   | False                            |

Table 2. Flexible controlling of the questionnaires in the IIA data collection module by using simple excel sheets
Furthermore, to analyze the cultural characteristics of the users, the value survey module (VSM94) has to be filled in by the user (cf. Hofstede, 2002). The VSM94 contains 26 questions to determine the values of the cultural dimensions using the indices of Hofstede that characterize the cultural behavior of the users (cf. Hofstede, 1991). The questions are implemented within the IIA data collection tool (cf. figure 11) as flexible questioning module which can be controlled by a simple excel sheet (cf. table 2): not only the contents but also the kind of questions can be defined (nominal, ordinal, interval, with/without qualitative reason/text box, with/without numerical display, checkboxes or radio buttons).

After this, the results of the VSM94 and those of the test tasks are presented to the user who has to estimate whether or not the cultural and informational values found correlate or match to him (cf. figure 12).

![Image](image_url)

Fig. 12. Asking the user to evaluate the results found during the IIA test

The debriefing part reveals the purpose of the test to the user in greater detail. It collects data regarding the usability of the test system, the perceived difficulty of the test in general,
if the user has recognized the implemented hypotheses in the test tasks during the test session, as well as e.g. asking the physical conditions of the test environment (cf. figure 13).

![Cross-Cultural Interaction Analysis](image)

Fig. 13. Asking the user about the conditions of the test environment

5. Data Collection and Data Analysis with the IIA Tool

Two online studies timely separated by one year (in 2006 and 2007) served to verify the functionality and reliability of the IIA tool and to get the preferences of users according to their cultural background (especially regarding their interaction behavior). Randomly selected employees from SiemensVDO (now Continental) all over the world were invited per email to do the test session using the IIA data collection module by downloading it from the corporate intranet. The test participant (Siemens VDO employee) downloaded the IIA data collection module via the corporate intranet locally on his computer, started the tool and did all test tasks. Before closing the tool, the collected data has been transferred automatically onto a non public and secure network drive on a SV server by the IIA tool (figure 14). Using the IIA data analysis module, the data could have been analyzed there.
Table 3 characterizes the two online studies regarding sample size, tests downloaded, tests aborted, valid test data sets, and return rate.

| Study | Sample size | Survey period | Number of downloaded tests | Tests Aborted [%] | Number of valid test data sets | Return Rate [%] |
|-------|-------------|---------------|---------------------------|------------------|-------------------------------|----------------|
| 1     | 600         | 12/14/05 - 01/14/06 | 166                      | 41,5             | 102                           | 16,6           |
| 2     | 14500       | 11/14/06 - 01/19/07 | 2803                     | 66,8             | 916                           | 6,3            |

Table 3. Characterization of the two online studies conducted with the IIA tool

The tests have been aborted due to the following reasons: download time too long\(^2\), no time to do the test now, test is not interesting or appealing. This type of qualitative data helped to optimize the testing equipment and to steer the direction of data analysis by asking the user for the reasons of his behavior during the test (e.g. by open questions using text boxes). Only complete and valid data sets have been analyzed using the IIA data analysis module and the statistic program SPSS (cf. Bortz & Döring, 2005). The discrimination rate of classifying the

\(^2\) Notably in China because of slow network connections.
users to their selected test language by the variables concerning the cultural background of
the user’s mother tongue, nationality, country of birth and primary residence was 83.3% for
the first and 81.9% for the second study. Therefore, the differences in HCI in these studies
have been analyzed in relation to three groups of test persons according to the selected test
languages (Chinese (C), German (G), and English (E)) in order to reduce data analyzing
costs. In the following, I concentrate on the more representative second main study, because
it has been used nine times more valid test data sets (916:102). Furthermore, the second
study almost mirrors the results of the first study. Nevertheless, I will contrast and discuss
the differing results in some detail to be able to deduce the reliability of the IIA tool. Out of
the 14500 test persons invited in the larger second main study, 2803 downloaded and started
the test. The return rate of 19.3% is sufficient for reasonable statistical analysis. 66.8% of
the tests have been aborted. The remaining 33.1% of the tests have been completed and only
the data of these tests has been analyzed using the IIA data analysis module and the statistics
program SPSS. The total remaining amount of valid data sets is 916.

To analyze the collected data, structural equation models have been used. Structural
equation models belong to the statistical methods of conformational factor analysis. In
contrast, explorative factor analysis can be used to determine the correctness of the conducted
classification of the parameters into factors (e.g. informational dimensions). Factor analysis
serves to structure and to select the deduction of the cultural interaction indicators (CII’s) of
the information dimensions for HMI design. The objective of factor analysis is the grouping
and reduction of the information quantity (judgments, questions, variables) simultaneously
ensuring and protecting of information content. The main tasks of factor analysis are:
• Grouping variables to „factors“ according to their correlation strength.
• Identifying variables resp. factors that correlate highly with information related
dimensions or predict them greatly.
• Filtering of variables having low explanation value in regard to the factor or the
informational dimension they represent.
• Resuming variables to indicators on the basis of factor and item analysis as well as
reflections regarding content.
• Deducting indicators representing (parts of) the information related dimensions.

The methods used for these purposes are explorative factor analysis, regression analysis and
item analysis (all feasible in SPSS). The analysis of the empirically collected data comparing
the average values using the IIA analysis module, neural networks, and AMOS revealed
that some of the parameters do really depend on culture.

6. Results: Cultural Interaction Indicators (CII’s) and Patterns (CIP’s)

In the two online studies, some values of the implemented variables in the IIA tool showed

3 The discrimination rate has been calculated using discrimination analysis (cross validated and
grouped, Wilk's Lambda in study 1: λ1,2=.072**, λ2=.568**, Wilk's Lambda in study 2: λ1,2=.192**, λ2=.513**). The level of significance is referenced with asterisks in this chapter (* p<.05, **
p<.01).
4 Cf. Backhaus et al., 2003.
5 AMOS is short for Analysis of MOment Structures. It is a statistical tool for data analysis providing
structural equation modeling (SEM). For further details, please refer to Arbuckle, 2005.
significant differences, which represent differences in user interaction according to the different cultural background of the users. Therefore, these variables can be called cultural interaction indicators. Table 4 presents the cultural interaction indicators that can be derived from the quantitative results of the two online studies.\(^6\)

| Cultural interaction indicator | First study | Second study |
|-------------------------------|-------------|--------------|
| MG.CarSpeed                   | F(2,102)=8,857** | \(\chi^2\) (2,916)=29,090** |
| MG.MessageDistance            | F(2,102)=7,645** | F(2,916)=16,241** |
| MD.NumberOfPOI                | F(2,102)=3,143* | \(\chi^2\) (2,916)=32,170** |
| MaximalOpenTasks              | \(\chi^2\) (2,102)=12,543** | F(2,916)=15,140** |
| MaximalOpenTasksRatio (C,G,E) | 2.5 : 1.4 : 1 | 1.7 : 1.03 : 1 |
| MG.InfoPresentationDuration   | \(\chi^2\) (2,102)=17,354** | \(\chi^2\) (2,916)=82,944** |
| NumberOfChars                 | \(\chi^2\) (2,102)=16,452** | \(\chi^2\) (2,916)=67,637** |

Table 4. Cultural Interaction Indicators found in both studies

The significant cultural interaction indicators are the following: MG.CarSpeed (\(\chi^2\) (2, 916) = 29,090**) means the driving speed of the simulated car in the maneuver guidance test task ((C) less than (G) and (E)). MG.MessageDistance (F (2, 916) = 16,241**) denotes the temporal distance of showing the maneuver advice messages in the maneuver guidance test task. (C) desired about 30% more pre-advices (“in x m turn right”) than (G) or (E) before turning right. This can be an indication for higher information speed and higher information density in China compared to Germany, for example. MD.NumberOfPOI (\(\chi^2\) (2, 916) = 32.17**) counts the number of points of interest (POI) set by the user in the map display test task. Information density increases with the number of POI and is two times higher for (C) than for (G) or (E). MaxOpenTasks (F (2, 916) = 15.140**) represents the maximum number of open tasks in the working environment (i.e. running applications and icons in the Windows™ task bar) during the test session with the IIA data collection module. (C) tend to work on more tasks simultaneously than (G) or (E) (ratio (C,G,E) = 1.7:1.03:1) which can be possibly explained by the way of work planning (polychrome vs. monochrome timing, cf. Hall 1976) or the kind of thinking (mono-causal (sequential) vs. multi-causal (parallel) logic, Röse et al., 2001). MG.InfoPresentationDuration represents the time the maneuver advice message is visible on the screen. (C) and (G) wanted the advices to be about 40% longer than (E) do. NumberOfChars (\(\chi^2\) (2) = 67.637**) contains the number of characters entered by the user during the maneuver guidance and map display test tasks in answering open questions ((C) < (E) and (G)). This is explained by the fact that the Chinese language needs considerably less characters to represent words than English or German.

There are also possible cultural interaction indicators that are only significant in the second study which is more representative than the first because of n=916 in comparison to n=102 (cf. table 5).

---

\(^6\) The variables in the valid test data sets are not distributed comparably in the first and the second online study. Therefore, partly the same variables have been analyzed either by ANOVA or by Kruskal-Wallis-test (indicated with F or \(\chi^2\)).
Variables with borderline values | First study | Second study
--- | --- | ---
OpenTaskBeforeTest | $F(2, 102) = 3.129^*$ | $\chi^2(2,916)=5,965$
OpenTaskBeforeTestRatio (C,G,E) | 1.6 : 1.2 : 1 | 1.05 : 1.04 : 1
IE.InteractionExactness | $F(2,102)=2,345 (p=.101)$ | $\chi^2(2,916)=24,106^{**}$
IS.InteractionSpeedValue | $F(2,102)=1,801 (p=.170)$ | $F(2,916)=16,246^{**}$
MG.NumberOfManeuver | $\chi^2(2,102)=4,785 (p=.091)$ | $\chi^2(2,916)=54,051^{**}$
UV.UncertaintyAvoidanceValue | $\chi^2(2,102)=5,297 (p=.071)$ | $\chi^2(2,916)=26,239^{**}$
IS.InteractionExactnessValue | $F(2,102)=2,698 (p=.073)$ | $\chi^2(2,916)=40,862^{**}$

Table 5. Cultural interaction indicators with borderline values in the studies

OpenTaskBeforeTest represents the number of open tasks in the working environment (i.e. running applications and icons in the Windows™ task bar) before the test session with the IIA data collection tool began. (C) tend to work on more tasks simultaneously than (G) or (E) which can be possibly explained by the way of work planning (polychrome vs. monochrome timing, Hall 1976) or the kind of thinking (mono-causal (sequential) vs. multi-causal (parallel) logic, Röse et al., 2001). IE.InteractionExactness ($\chi^2(2, 916) = 24.106^{**}$) measures the exactness clicking onto dots in the abstract test task of “clicking dots away”. (G) clicked the dots away almost twice as exact as (E) and (C). IS.InteractionSpeedValue ($F(2, 916) = 16.246^{**}$) measures the duration of the abstract test task of “clicking dots away”. (C) clicked the dots away almost twice as fast as (E) and (G).

The number of mouse clicks differs in both studies but in different significance peculiarities: E.g., UV.MouseClicks counts the mouse clicks in the test task “uncertainty avoidance”. (C) are doing more than (G) and (E) which indicate the desire of (C) to get immediate system reaction according to their input requests (e.g. mouse clicks) is very high. This is also supported by AllMouseClicks ($\chi^2(2, 916) = 15.235^{**}$) which counts all mouse clicks done by a user in the test and the nearly twofold amount of ErrorClicks by (C) in contrast to (G) and (E). ErrorClicks ($\chi^2(2, 916) = 9.771^{**}$) counts the mouse clicks, which do not have any function for a test task (and hence, which can be a cue for impatience).

The peculations of the cultural interaction indicators regarding the different cultures are similar comparing them between first (n=102) and second (n=916) data collection. This indicates the correctness of the data collection. The analysis of the log files of the second data collection using 1632 valid data sets revealed the results shown in table 6.

Table 6. Cultural interaction indicators derived from log files of the second data collection

| df = 2 | Name of CII | Oneway ANOVA | Kruskal-Wallis | Interpretation |
| --- | --- | --- | --- | --- |
| **| ** | F | P | h | $\chi^2$ | p | CII is significant |
| Test Duration | 11,53 | 0,000 | 0,404 | 54,508 | 0,000 | yes, quantitatively |
| MouseMoves_norm | 26,20 | 0,000 | 0,225 | 57,900 | 0,000 | yes, quantitatively |
| KeyDowns_norm | 27,31 | 0,000 | 0,318 | 59,451 | 0,000 | yes, quantitatively |
| LeftButtonDowns_norm | 28,84 | 0,000 | 0,266 | 59,471 | 0,000 | yes, quantitatively |

It is remarkable that all three CII’s (MouseMoves_norm, KeyDowns_norm, and LButtonDown_norm) concerning the kind of interaction behavior of the users are peculiarized very similar according to the nationality of the test participants indicating the same interaction behavior of the users of the same nationality which indicates the correctness of the test equipment and the study results.

The cultural interaction indicators can be visualized applying the IIA data analysis tool to plot “cultural HCI fingerprints” (in the style of Smith & Chang, 2003) which represent the cultural differences in HCI in respect to several variables for HCI design that depend on the cultural background of the potential target group of users (cf. figure 15). This visual representation of the CII’s should ease information reception and improve comparative understanding of the cultural differences in HCI.

![Figure 15. "Cultural HCI Fingerprints" (different values of the cultural interaction indicators according to test languages) plot by the IIA data analysis tool](image)

7. Discussion: Reliability of Results, IIA Tool and Design Recommendations

The two main online studies in this work revealed many aspects, which supported each other: a high discrimination rate of over 80% and the high accordance between the cultural interaction indicators found by one-way ANOVA and Kruskal-Wallis-Test respectively and
the discrimination analysis on the other hand supports the high reliability and criteria validity of the statistical results received in this study using the IIA tool. Moreover, the tests with the IIA data evaluation tool using neuronal networks confirmed also the high classification rates of the several combinations of cultural interaction indicators of over 80%. This overall outcome proofs the high reliability and justifies the usage of the IIA tool in future. Furthermore, the reliability of the results and the IIA tool is also supported by the fact, that the results of other qualitative studies confirmed the results of the quantitative studies done with the IIA tool: the studies with the IIA tool comparing Chinese and German users revealed different interaction patterns according to the cultural background of the users regarding e.g. design (ample vs. simple), information density (high vs. low), menu structure (high breath vs. high depth), personalization (high vs. low), language (symbols vs. characters) and interaction devices (no help vs. help) that have been confirmed by qualitative studies e.g. by Vöhringer-Kuhnt 2006, Kralisch 2006 or Kamentz 2006. Results regarding e.g. the status bar position of the URD test task are qualitatively confirmed by e.g. Röse 2001. The quantitative studies gave some first insights into the possibility of classifying cultural different users on behalf of their interaction behavior concerning the direct hidden cultural variables. The results are reliable because they have been traceable and reproducible by the two online studies, finally yet importantly because of the high sample size. However, detailed values and higher discrimination power of the CII’s have to be determined in future. Additionally, qualitative studies brought to light some results concerning the direct visible cultural variables. However, the sample size of the qualitative studies done was very small and hence, these results can only give direction to new guidelines instead of being precise guidelines for the future. A critical general objection to the application of the results yielded by the two studies using the IIA tool could concern the fact, that the collected data are selective samples because they are restricted to use cases of driver navigation systems and to employees of SiemensVDO. Hence, it is not allowed to generalize the results for all Chinese and all German users. Nevertheless, it is permitted and necessarily indicated to extract thumb rules for intercultural HMI design, because the results of the studies revealed that there is a metrics, which is adequate to measure cross-cultural HCI. Although the VSM values are similar to Chinese, German, and English speaking employees of SV (probably because of their common company philosophy) and their experience in working with computers is alike, the HCI between Chinese, German, and English speaking employees of SV differs significantly. Hence, some results can be expected to be valid for HCI design in general because there are culturally sensitive variables that can be used to measure cultural differences in HCI only by counting certain interaction events without the necessity of knowing the semantic relations to the application. Such indicators are e.g. mouse moves, breaks in the mouse movements, speed of mouse movements, mouse clicks and interaction breaks. Surely, all those indicators can also be connected semantically to the use cases or applications. However, simply counting such events related to the session duration from users of one culture and comparing them to users of another culture is obviously sufficient to indicate differences in interaction behavior of culturally different users. The possible implication that this is grounded in subconscious cultural differences imprinted by primary culture and learning the mother tongue which leads to different HCI independently of the conscious cultural propositional attitudes, has to be verified in future studies. Additionally, studies that are more detailed must show whether changing the metrics of potential cultural indicators (or using them in other situations, use cases, or circumstances) will improve their
discriminating effect and yield appropriate values accordingly as well as the general usability of general cultural interaction indicators. The following criteria, which represent the real UMTM, have been identified as hypothetically depending on culture by literature research and reflection (according to the IV model) as well as actually depending on culture by empirical research (using the IIA tool):

- Information frequency (number of words per minute, number of dialogs per minute)
- Information density (number of images per page, number of words per dialog, number of words per information unit, image-text ratio or distribution, distance of pieces of information to each other)
- Information arrangement (widget positions, image-text arrangement)
- Information order (regularity and orderedness of informational units)
- Information sequentiality (sequential presentation of information units)
- Interaction speed (mouse clicks per minute, overall mouse clicks, length of mouse track per second)
- Interaction device (mouse, keyboard, menu control button, touch screen)

However, these confirmations of the postulated hypotheses do not finally proof cultural differences in the information related dimensions. One important reason for this is that the used metrics of the test setting must be optimized regarding the use cases and their logical relevance. Furthermore, the reliability of the metrics and the used indicators has to be determined more exactly and optimized. This requires applying test theory in much more detail.

8. Conclusion and Outlook

The IIA tool serves to record the user’s interaction with the computer to be able to identify cultural variables like color, positioning, information density, interaction speed, interaction patterns and their values, which enable the deduction of design rules of thumb for cross-cultural HCI design. It is effective, efficient and reasonable to use the IIA tool within the process of cross-cultural HCI design because it can be used locally and worldwide and provides quantitative comparable and reliable results whose validity and method to get them is quantitatively and qualitatively confirmed by several studies. The results of two longitudinal studies proofed the reliability of the IIA tool and indicated that it should be possible to optimize the model of cultural dependent variables for the HMI design using structure equal models. Using the IIA tool means rapid use case design, i.e. real-time prototyping of user interfaces for different cultures as well as collecting huge amounts of valid data rapidly and easily worldwide online via internet or intranet.

The IIA tool will be continually optimized based on user feedback to extend the analysis and evaluation of cultural differences in HCI by exploring cultural interaction patterns and by improving the discrimination capability of cultural interaction indicators. Questionnaires in conjunction with recording biofeedback signals (heart rate and skin response) should give controlled insights into the user preferences. Another objective is to develop enhanced techniques using statistical methods (factors analysis, structure equation models, cluster analysis etc.), data mining and semantic processing to extract the cultural variables and its values as well as the guidelines for cross-cultural HMI design in a more automatically way. Moreover, it is intended to extend the method to implement new use cases e.g. by employing authoring tools or using HMI description languages.
9. References

Arbuckle, J. L. (2005). Amos 6.0 User’s Guide. Amos Development Corporation. Spring House. USA

Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2003). Multivariate Analysemethoden: Eine anwendungsorientierte Einführung. Springer. Berlin, Heidelberg, New York. 10. Aufl.

Bortz, J. & Döring, N. (2005). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3. Aufl. Springer. Heidelberg

De la Cruz, T., Mandl, T., Womser-Hacker, C. (2005). Cultural Dependency of Quality Perception and Web Page Evaluation Guidelines: Results from a Survey. In Day, D., del Galdo, E., Evers, V. (eds.): Designing for Global Markets 7: Proceedings of the Seventh International Workshop on Internationalization of Products and Systems IWIPS 2005, The Netherlands, Grafisch Centrum, Amsterdam, 15–27

Dix, A., Finlay, J., Abowd, G., Beale, R. (1993). Human-Computer Interaction. Prentice Hall, London

Gellner, M. & Forbrig, F. (2003). ObSys – a Tool for Visualizing Usability Evaluation Patterns with Mousemaps. HCI International, Greece

Hall, E.T. (1959). The Silent Language. Doubleday, New York

Hall, E.T. (1976). Beyond Culture. Anchor Press, Garden City, N.Y.

Halpin, A.W. & Winer, B.J. (1957). A factorial study of the leader behavior descriptions. Bureau of Business Research, Ohio State University. Columbus, Ohio

Hamacher, N. (2006). Automatic Criteria Oriented Usability Evaluation of Interactive Systems. Dissertation. München: Dr. Hut-Verlag

Heimgärtner, R. (2005). Research in Progress: Towards Cross-Cultural Adaptive Human-Machine-Interaction in Automotive Navigation Systems. In: Day, D., del Galdo, E.M. (eds.): Proceedings of the Seventh International Workshop on Internationalization of Products and Systems. IWIPS 2005, The Netherlands, Grafisch Centrum, Amsterdam, 7–111

Heimgärtner, R. (2006). Measuring Cultural Differences in Human Computer Interaction as Preparatory Work for Cross-Cultural Adaptability in Navigation Systems. In: Useware 2006, VDI-Bericht Nr. 1946, VDI-Verlag, Düsseldorf, 301–314

Heimgärtner, R. (2007). Cultural Differences in Human Computer Interaction: Results from Two Online Surveys. In: Osswald, A., Stempfhuber, M., Wolff, C. (eds.). Open Innovation. Proc. 10. International Symposium for Information Science. Constance: UVK. 145-157

Herczeg, M. (2005). Software-Ergonomie. Grundlagen der Mensch-Computer-Kommunikation. 2. Auflage, München: Oldenbourg.

Hofstede, G. (1991). Cultures and Organisations: Software of the Mind. McGraw-Hill, London

Hofstede, G. (1994). Value Survey Module 1994 Manual, Institute for Research on Intercultural Cooperation, Maastricht

Hofstede, G. (2002). The Pitfalls of Cross-National Survey Research: A reply to the Article by Spector et al. on the Psychometric Properties of the Hofstede Values Survey
Module 1994. Applied Psychology: An International Review, Vol. 51, Nr. 1 (2002) 170-173

Hofstede, G. & Hofstede, J. G. (2005). Cultures and Organizations: Software of the Mind. 2nd Edition. USA, New York: McGraw-Hill

Honold, P. (2000). Interkulturelles Usability Engineering: Eine Untersuchung zu kulturellen Einflüssen auf die Gestaltung und Nutzung technischer Produkte. Reihe 10: Informatik/Kommunikation. Düsseldorf: VDI Verlag GmbH

Kamentz, E. (2006). Adaptivität von hypermedialen Lernsystemen: Ein Vorgehensmodell für die Konzeption einer Benutzermodellierungskomponente unter Berücksichtigung kulturbedingter Benutzereigenschaften. Dissertation

Kamentz, E. & Mandl, T. (2003). Culture and E-Learning: Automatic Detection of a Users’ Culture from Survey Data. In Evers, Vanessa and Röse, Kerstin and Honold, Pia and Coronado, Jose and Day, Donald, Eds. Proceedings Fifth International Workshop on Internationalization of Products and Systems (IWIPS 2003) Designing for Global Markets 5: Proceedings of the Fifth International Workshop on Internationalization of Products and Systems (IWIPS 2003), pp. 227-239, Berlin, Germany

Kralisch, A. (2006). The impact of culture and language on the use of the internet – empirical analyses of behaviour and attitudes. Dissertation

Lewandowitz, L., Rößger, P., Vöhringer-Kuhnt, T. (2006). Asiatische vs. europäische HMI Lösungen von Fahrerinformationssystemen. In: USEWARE 2006 (VDI-Bericht 1946). VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik. Düsseldorf: VDI-Verlag, 301-314

Mangold, P. (2005). Proceeding studies on behavior - not only a challenge for professional tools. Usability Symposium 2005: 127-140

Marcus, A. & Baumgartner, V-J. (2005). A Practical Set of Culture Dimensions for Global User-Interface Development. In: Masoodian, M., Jones, S., Rogers, B. (eds.): Proceedings of Computer Human Interaction: 6th Asia Pacific Conference. APCHI 2004, New Zealand, Rotorua, 29 June – 2 July 2004, Volume 3101 / 2004. Springer, Berlin (2005) 252-261

Prendinger, H. & Ishizuka, M. (Eds.) (2004). Life-Like Characters. Springer

Röse, K. (2002). Methodik zur Gestaltung interkultureller Mensch-Maschine-Systeme in der Produktionstechnik. Fortschritt-Berichte pak. Band 5. Mensch-Maschine-Interaktion, Verlag Universität Kaiserslautern, Kaiserslautern

Röse K, Liu, L., Zühlke, D. (2001). Design Issues in Mainland China: Demands for a Localized Human-Machine-Interaction Design. In: Johannsen, G. (ed.): 8th IFAC/IFIPS/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems. Preprints, Kassel (2001) 17-22

Shneiderman, B. (2002). User Interface Design. MIT Press

Smith, A. & Chang, Y. (2003). Quantifying Hofstede and Developing Cultural Fingerprints for Website Acceptability. In: Evers, V., Röse, K., Honold, P., Coronado, J., Day, D. (eds.): Proceedings of the Fifth International Workshop on Internationalization of Products and Systems. IWIPS 2003, Germany, Berlin, 17-19 July 2003. University of Kaiserslautern, Kaiserslautern, 89-102

Victor, D. (1997). International Business Communication. New York, Prentice Hall
Vöhringer-Kuhnt, T. (2006). Asiatische vs. europäische HMI Lösungen von Fahrerinformationssystemen. In: Useware 2006, VDI-Bericht Nr. 1946, VDI-Verlag, Düsseldorf (2006) 279–287
The book consists of 20 chapters, each addressing a certain aspect of human-computer interaction. Each chapter gives the reader background information on a subject and proposes an original solution. This should serve as a valuable tool for professionals in this interdisciplinary field. Hopefully, readers will contribute their own discoveries and improvements, innovative ideas and concepts, as well as novel applications and business models related to the field of human-computer interaction. It is our wish that the reader consider not only what our authors have written and the experimentation they have described, but also the examples they have set.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ruediger Heimgaertner (2008). A Tool for Getting Cultural Differences in HCI, Human Computer Interaction: New Developments, Kikuo Asai (Ed.), ISBN: 978-953-7619-14-5, InTech, Available from: http://www.intechopen.com/books/human_computer_interaction_new_developments/a_tool_for_getting_cultural_differences_in HCI