Understanding Worldwide Human Information Needs

Revealing Cultural Influences in HCI by Analyzing Big Data in Interactions

Rüdiger Heimgärtner

HCI Research & Development, Intercultural User Interface Consulting (IUIC), Lindenstraße 9, 93152 Undorf, Germany

Keywords: Culture, Relation, Structure Equation Model, Neural Networks, Factor Analysis, Data Analysis, Global, Local, Intercultural Interaction Analysis, Tools, Big Data, Modeling, Data Mining, Statistics.

Abstract: Understanding human information needs worldwide requires the analysis of much data and adequate statistical analysis methods. Factor Analysis and Structural Equation Models (SEM) are a means to reveal structures in data. Data from empirical studies found in literature regarding cultural human computer interaction (HCI) was analyzed using these methods to develop a model of culturally influenced HCI. There are significant differences in HCI style depending on the cultural imprint of the user. Having knowledge about the relationship between culture and HCI using this model, the local human information needs can be predicted for a worldwide scope.

1 INTRODUCTION

Rapidly progressing globalization requires consequent adaptation of the methods and processes applied in Human Computer Interaction (HCI) design to the relevant cultural needs. Intercultural User Interface Design (IUID) integrates several disciplines. For example it integrates information technology with cultural studies within this prominent and complex field. However, the relations between HCI and culture are not yet well elaborated, even if there are some initial approaches (cf. e.g., Röse et al., 2001); (Marcus, 2006); (Vatrapu and Suthers, 2007); (Clemmensen, 2009); (Heimgärtner, 2012). A proper taxonomy of all approaches, methods and processes in IUID is also still missing, even if there are some clues (cf. e.g., Clemmensen and Röse, 2010). In addition, there are several tools for user interaction logging (e.g. ObSys for recording and visualization of windows messages, cf. Gellner and Forbrig, 2003). However, none of the existing tools provide explicitly cultural usability metrics to measure culturally imprinted interaction behavior. This is because the connections between HCI and culture are not systematically collected and prepared for such tools.

This paper describes some methods to elucidate the connection between HCI and culture by analyzing big data (cf. Auinger et al., 2011): the intercultural interaction analysis tool (IIA tool), neural networks, factor analysis and structural equation models (SEM). Then, the application of the methods to analyze the connections between culture and HCI are explained and the results and challenges are addressed.

2 METHODS FOR ANALYZING CULTURE IN HCI

The following methods and tools can be used to yield cultural differences in HCI and to determine the relationship between culture and HCI. Cultures are orientation systems for group members (cf. Thomas et al., 2010). The characteristics of cultures can be described using cultural dimensions (cf. Hofstede and Hofstede, 2005). The interaction behavior of the user with the system can be described with HCI dimensions (cf. Heimgärtner 2012). The aim is to determine the connection between cultural dimensions and HCI dimensions and its values using adequate methods and tools to yield a model for culturally influenced HCI that serves to predict the HCI style of members of any cultures. With this knowledge, relevant design recommendation can be derived to develop user interfaces with high usability. Quantitative values for the indices of the cultural dimensions are available from empirical studies by Hofstede and Hofstede 2005. Further quantitative values for the

Heimgärtner R..
Understanding Worldwide Human Information Needs - Revealing Cultural Influences in HCI by Analyzing Big Data in Interactions.
DOI: 10.5220/0004166401650169
In *Proceedings of the International Conference on Data Technologies and Applications* (DATA-2012), pages 165-169
ISBN: 978-989-8565-18-1

indicators of the HCI dimensions are also available (cf. Heimgärtner, 2012). 916 valid data sets together with the values of the indices of Hofstede are used to determine the connections between the values of the cultural and HCI dimensions.

2.1 IIA Tool as a Method Framework

A special tool for measuring the interaction behavior of the user has been developed by the author (cf. Heimgärtner, 2008). The interaction of culturally different users doing the same test task can be observed (using the same test conditions i.e. the same hard- and software, environment conditions, language and the test tasks) as well as requiring the same experience with the use of the system. Logging data of dialogs, debugging and HCI event triggering while using the system are highly valuable (Kralisch, 2006). This data can be logged during usability tests according to certain user tasks. The IIA tool provides data collection, data analysis and data evaluation. It serves to record and analyze the user's interaction with the system in order to identify culturally influenced variables such as color, positioning, information density and interaction speed as well as their values. The collection and preparation of the data is to be done automatically for the most part by the IIA data collection module. The data is to be stored in databases in a format that is immediately usable by the IIA data analysis module, which does subsequent data conversion or preparation. Common statistic programs like SPSS or AMOS can be deployed to apply statistical methods (cf. Ho, 2006). The IIA data evaluation module enables classification using neural networks to cross-validate the results from data analysis.

The Delphi-IDE allows transformation of new HCI concepts very quickly into well formed prototypes that can be tested very soon within the development process. For example, some hypotheses have been confirmed quantitatively addressing many test users online in 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 as well as a very large amount of valid data collected quickly and easily worldwide online via internet or intranet or offline locally on the spot.

2.2 Using Neural Networks

Neural networks are used within the IIA evaluation module to verify and establish trends of cultural differences in user interaction and to enhance the plausibility of quantitative results. For example, it might not be important which subjects take part in a test, if neural networks are used, which can independently learn existing trends (e.g., back propagation networks (cf. Haykin, 2008) or self-organizing maps (cf. Kohonen, 2001)). Therefore, such networks are not concerned about which test persons take part in the test. By connecting the categorized grouped test data according to the HCI dimensions to the input neurons and the cultural characteristics (represented for example by variables like nationality, mother tongue, etc.) of the users to the output neurons of the neural network, training of the network will reveal if there is a correlation of the values of the HCI dimensions with the input and the culture with the output of the neural network. In other words, if cultural differences do exist, i.e. if there is a correlation between the corresponding test data of the test persons and the culture at the output of the neural network, the neural network will learn and reveal it ("monitored learning", cf. Mandl et al., 2003). By means of connecting test data which is categorized or grouped according to hypotheses and the cultural variables to the output of the neural network, it identifies whether cultural differences do exist and to which degree. Thereby, it can be seen whether correlations exist between the test data from the subject with the input of the neural network and his cultural imprint with the output of the neural network.

2.3 Explorative Factor Analysis

HCI dimensions represent classes of indicators. For example, the class "number of *information units* per *space* unit" belongs to the HCI dimension "information density" and can be expressed by the indicator "number of words displayed on a screen". Another HCI dimension is "interaction frequency". This dimension contains the class "number of *interactions* per *time* unit" (e.g., represented by the indicator "number of mouse clicks per second"). Support for the correctness of the HCI dimensions comes from the application of factor analysis methods.

Explorative factor analysis serves hereby to derive the HCI dimensions by grouping potential variables to factors (component matrix) according to the strength of their correlation (correlation matrix) (cf. Ho, 2006). Using main component analysis with the 916 data sets including 19 indicators, four main components have been extracted (ranked by impact, cf. Heimgärtner, 2012): (i) Style of user interaction behavior (expressed by interaction speed index and interaction exactness index) is influenced by age, time abroad and experience with the use of computers as well as task orientation of the user. (ii) Information reception (represented by information density, number and order index) varies because of the influence of uncertainty avoidance. (iii) Information speed (represented by the information speed index) correlates with help usage (represented by the number of help calls). Furthermore, (iv) gender and power distance indices seem to be correlated. The four main factor loadings ("components") derived from the data explain nearly 95% of the variance. The remaining 15 components in total explain only 5% of the variance. This analysis is clear, if not very meaningful because the extensions of the indicators constituting the components partly overlap. Therefore, an "extended" (more detailed) factor analysis has been carried out. It revealed that component #1 (interaction behavior) in the "simple" factor analysis can be split up into the two components "primary cultural imprint" and "interactional behavior" or "HCI style". Hence, the largest components concern HCI and cultural aspects, which now can be further used to analyze their connections doing structural equation modeling.

2.4 Structural Equation Modeling

Structural equation models (SEM) belong to the statistical methods of confirmative factor analysis, which can be performed using e.g., AMOS (cf. Byrne, 2001). A structural equation model consists of a set of equations. The effect or endogenous variable is on the left side and on the right side is the sum of the causes with each causal variable multiplied by a causal parameter. Here, structural equation modeling serves to identify the relationships between cultural dimensions and HCI dimensions and their causes. SEM has, on the left hand side the cultural indices by Hofstede representing cultural dimensions and on the right hand side, the indicators of the HCI dimensions. Cultural dimensions and HCI dimensions are connected by assumed relations in the SEM. The theory is the better the more variances in the empirical data can be explained statistically by the SEM. The theory depending on the modeled relations is best if the balance between the variables is in equilibrium. Adding or removing variables or relations to improve the equilibrium is called structural equation modeling. First attempts using AMOS showed that the higher the relationship orientation (collectivism) represented by Hofstede's individualism index (IDV), the higher the information density, information speed, information frequency, interaction frequency and interaction speed as well as the other way around. Additional methods like conformational factor analysis as well as regression analysis support this process of finding the right model.

3 REVEALING CULTURAL INFLUENCES IN HCI

3.1 Results

The data analyzed for the qualitative and quantitative studies revealed a trend for the investigated cultures that allowed shifting towards a model of culturally influenced HCI. (cf. Heimgärtner, 2012). With the right combination of indicators representing the HCI dimensions, it is possible to capture interaction differences that are culturally imprinted (e.g., according to cultural aspects such as nationality, mother tongue, country of birth, etc. or to cultural dimensions). There are correlations between the interaction of the users with the system and their cultural background. The cultural differences in HCI concern layout (complex vs. simple), information density (from high to low), personalization (from greater to lesser), language (symbols vs. characters), interaction speed (from high to low) and interaction frequency (from high to low). In addition, the cultural differences found in HCI by discriminance analysis are quantitatively measurable by a computer system using a special combination of indicators represented by interaction patterns depending on the culturally imprinted interaction behavior of the user. The recognition and classification of cultural inter-action patterns in HCI, i.e. cultural differences in HCI, can be achieved purely *quantitatively* (cf. Heimgärtner, 2012). A handful of indicators are sufficient for this purpose. Moreover, the interaction patterns representing the cultural differences in HCI and the derived indicators are statistically sufficiently discriminating to enable computer systems to detect them *automatically* and to assign the users to a certain cultural imprint. Furthermore, when reversed, interaction patterns are also useful to identify a user, which is necessary for the system to adapt to the corresponding user.

3.2 Challenges

However, many open issues still remain. The reliability of the test equipment and test methods as well as the results and models must be discussed and alternative test settings in the future should be used. The results have been determined statistically and are mainly descriptive. Explaining the inference statistics for the evidence and the reasons of the cultural differences in depth remains to be done. In addition, there is the need for strengthening the confirmation of the HCI dimensions by conducting deeper explorative factor analysis with the data from further studies as well as enhancing the separation effect and discriminatory power of the indicators and their classes. The test data sets must be evaluated in more detail to generate optimized algorithms for cultural adaptability in HCI based on neural networks, which need large amounts of interaction data for training, validating and testing as well as on structured equal models to prove basic theoretical and well explained interaction models by taking cultural aspects into account.

4 CONCLUSIONS

Many kinds of culturally influenced interaction patterns are only recognizable over time requiring the collection of big data. Hence, enhanced algorithms and tools must be used for data analysis. Therefore, designing tools for non-experts to do their own analysis with big data in interactions will be a prominent task for interaction designers in the future (cf. Fisher et al., 2012). The combination of different statistical methods to determine cultural differences and influences in HCI represents an initial idea and the first step to ascertain the right relationships between culture and HCI. However, much effort still remains. Nevertheless, the presented approach and model are worthy of being investigated and optimized in the future. Revealing cultural influences in HCI by analyzing big data in interactions to finally create and use a model or even a theory for culturally influenced HCI should help to better understand human information needs worldwide.

REFERENCES

Auinger, A., Aistleithner, A., Kindermann, H., Holzinger, A., 2011. Conformity with User Expectations on the Web: Are There Cultural Differences for Design Principles? In: Marcus, A. (Ed.) *Design, User Experience, and Usability. Theory, Methods, Tools and Practice. Lecture Notes in Computer Science LNCS 6769/2011*. Springer. Heidelberg. Berlin. New York. 3-12.

Byrne, B. M., 2001. *Structural equation modeling with AMOS: Basic concepts, applications, and programming,* Lawrence Erlbaum. Mahwah, NJ.

Clemmensen, T., 2009. Towards a Theory of Cultural Usability: A Comparison of ADA and CMU Theory. In: *HCD 2009, Held as Part of HCI International 2009*, San Diego, CA, USA, July 19-24.

Clemmensen, T., Röse, K., 2010. An Overview of a Decade of Journal Publications about Culture and Human-Computer Interaction (HCI). In: *Human Work Interaction Design: Usability in Social, Cultural and Organization Contexts. IFIP Advances in Information and Communication Technology*, Volume 316/2010, 98-112.

Fisher, D., DeLine, R., Czerwinski, M., Drucker, S., 2012. Interactions with big data analytics. In: *Interactions*, Volume 19, 3 (May 2012). ACM. 50-59.

Gellner, M., Forbrig, P., 2003. ObSys–a Tool for Visualizing Usability Evaluation Patterns with Mousemaps. In: Jacko, J., Stephanidis, C., (Eds.), *Human-Computer Interaction, Vol. I, Theory and Practice, Proceedings of the 10th International Conference on HCI*. Lawrence Erlbaum Associates. Mahwah, New Jersey, London. 469-473.

Haykin, S., 2008. N*eural networks: a comprehensive foundation*, Prentice Hall.

Heimgärtner, R., 2008. A Tool for Getting Cultural Differences in HCI. In: K. Asai, *Human Computer Interaction: New Developments*. InTech. Rijeka. 343-368.

Heimgärtner, R., 2012. *Cultural Differences in Human Computer Interaction - Towards Culturally Adaptive Human Machine Interaction*. Dissertation. University of Regensburg. Oldenbourg. München.

Ho, R., 2006. *Handbook of univariate and multivariate data analysis and interpretation with SPSS*, Chapman & Hall/CRC. Boca Raton, Fla.

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

Kohonen, T., 2001. *Self-organizing maps*, Springer. Berlin.

Kralisch, A., 2006. *The Impact of Culture and Language on the Use of the Internet Empirical Analyses of Behaviour and Attitudes*, Dissertation. Berlin.

Mandl, T., Schudnagis, M., et al., 2003. A Framework for Dynamic Adaptation in Information Systems. In: *C. Stephandis. Human Computer Interaction: Theory and Practice. Proceedings of 10th International Conference on Human-Computer Interaction*, 22-27 June 2003. Crete. 425-429.

Marcus, A., 2006. Cross-Cultural User-Experience Design. In: *Diagrammatic Representation and Inference.*16-24.

Röse, K., Zühlke, D., et al. 2001. Similarities and

Dissimilarities of German and Chinese Users. In: G. Johannsen. *Preprints of 8th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*. Germany. Kassel. 24-29.

Thomas, A., Kinast, E.-U., Schroll-Machl, S., 2010 (Eds.). *Handbook of intercultural communication and cooperation. Basics and areas of application*. Vandenhoeck & Ruprecht. Göttingen.

Vatrapu, R., Suthers, D., 2007. Culture and Computers: A review of the concept of culture and implications for intercultural collaborative online learning. In: *Lecture Notes in Computer Science 4568*. Springer. 260-275.