A cautionary tale: Hofstede's VSM revisited

Lidia Oshlyansky
University of Wales Swansea
Swansea, SA2 8PP
lidiaosh@gmail.com

Paul Cairns
UCL Interaction Centre
London WC1E 7DP
p.cairns@ucl.ac.uk

Harold Thimbleby
University of Wales Swansea
Swansea, SA2 8PP
h.w.thimbleby@swansea.ac.uk

Hofstede's cultural model has been widely used to make sense of the differences seen in cross-cultural HCI research.Hofstede's Value Survey Module (VSM) and the cultural indices it produces are well known in the HCI community. This paper reports on a recent re-examination of the VSM (specifically, VSM94) in nine countries. Contrary to expectations, the data collected did not replicate Hofstede's findings. Subsequent factor analysis reveals that VSM questions are not resulting in robust, replicable factors. We discuss possible issues in the method of data collection, but given that our method follows that of many other similar studies, our results suggest that the VSM should be interpreted with caution, particularly if it is to be used to adapt interfaces for different cultures.

VSM, Culture, Cross-cultural HCI, Cultural indices

1. HOFSTEDE'S VSM IN HCI

The field of HCI had investigated the effects of culture on user interface design, use, adoption and usability of various technologies. Often culture is studied within the framework of Hofstede's classic cultural model [1, 2], which is based on the five dimensions of the Value Survey Module (VSM). The VSM dimensions have been widely used as a framework to provide design guidelines, for instance for cross-cultural website development [3]. Other researchers have attempted to test guidelines created by Marcus by matching subjects' cultural profile to the cultural profile of the website, though with mixed results [4]. Hofstede's model has also been used as a framework to explain differences in existing website designs [5]. The Hoffsted model is widely cited; indeed, out of 6 papers in the 2005 British HCI conference discussing culture, 4 cited Hofstede.

Despite its wide use and acceptance, Hofstede's work has been criticised for its lack of insight into the richness and depth of culture. It has been suggested that a more qualitative or activity theory approach would be more appropriate [6, 7]. Hofstede's research has also been criticised because it focuses on national cultures. The study's methodology has also been brought into question as, for example, his entire sample was drawn from IBM employees [8].

Nonetheless, given the extensive use of the Hofstede model, it may still have value in capturing some aspects of usability cross-culturally. We performed a study to put the implications of the VSM on a quantitative footing by studying it in relation to other commonly used instruments in HCI. Our approach fits with Hofstede's view of using the VSM to conduct secondary research [9]. This paper focuses solely on the data collected from the VSM; a larger study and analysis is still underway. Surprisingly, the new data does not replicate any of the rankings of countries against the different dimensions found in Hofstede's previous work. Principal component analysis was used to analyse our VSM data in more detail and none of Hofstede's original dimensions are reflected robustly in the new data. As the method of our study follows that of Hofstede's and others in using the VSM to measure culture, our results suggest that the cultural dimensions are not as widely generalizable as previous literature would suggest.

2. METHOD OF DATA COLLECTION

The aim of the study was to collect VSM data (amongst others) from students in eleven countries,which would replicate Hofstede's work and also update it with a younger, different demographic than the IBM employees originally used. Hofstede's VSM94 was translated into six languages: Arabic (Saudi Arabian) Czech, Dutch, French, Greek and Malay. All translations were done by two bilingual speakers using the back-translation process to ensure as much accuracy as possible [10]. The questionnaires were distributed to university students in the Czech Republic, France, Greece, India, Malaysia, Netherlands, New Zealand, Saudi Arabia, South Africa, the United Kingdom, and to the United States.

A total of 1428 questionnaires were returned. Countries were included in the present analysis only if close to 100, or more, questionnaires were available for analysis from that country, to ensure an appropriately representative dataset from each country. Only those questionnaires that were completed by natives of the country being sampled were used in the analysis – native" being defined as having been wholly educated in and living in the country being sampled. Unfortunately, despite our best efforts, insufficient responses were received from France and the Netherlands, leaving nine countries: Czech Republic, Greece, India, Malaysia, New Zealand, Saudi Arabia, South Africa, the United Kingdom and the United States. The highest number of questionnaires was returned from Malaysia (168) and the lowest from Saudi Arabia (91). The final sample consisted of 519 men and 489 women (72 respondents did not give their gender). The average age was 23.4 years.

3. VSM SCORES AND RANKINGS

For each of the nine countries, the scores for each dimension were calculated using the formulas provided by Hofstede [9]. These scores are shown in Table 1. In brackets below each score is the ranking for the given dimension obtained in our study, followed by the ranking based on Hofstede's scores. Saudi Arabia has no rankings as it was not studied as an individual country by Hofstede; in addition there is only limited information available on Time Orientation.

	Czech Republic	Greece	India	Malaysia	New Zealand	United States	South Africa	United Kingdom	Saudi Arabia	Kendall rank correlation
Power Distance	35.0 (2/4)	50.4 (1/3)	31.4 (3/2)	23.4 (5/1)	20.9 (7/8)	20.8 (8/6)	23.2 (6/5)	30.2 (4/7)	29.1	0.38
Uncertainty Avoidance	83.2 (7/2)	113.4 (1/1)	97.0 (3/6)	97.4 (2/7)	81.3 (8/3)	83.9 (5/5)	89.3 (4/3)	83.7 (6/8)	93.6	-0.36
Individualism	85.1 (6/5)	94.0 (4/7)	78.0 (8/6)	80.4 (7/8)	96.4 (3/3)	97.3 (2/1)	87.3 (5/4)	103.3 (1/2)	88.2	0.64
Masculinity	17.0 (6/5)	45.2 (2/5)	49..0 (1/7)	33.0 (4/8)	11.2 (7/4)	31.8 (5/3)	34.7 (3/2)	7.3 (8/1)	49.9	-.40
Time Orientation	54.0 (4/5)	56.2 (1/)	42.2 (8/1)	54.3 (3/)	51.6 (5/2)	46.2 (7/3)	48.3 (6/)	54.4 (2/4)	42.4	-0.60

TABLE 1: VSM Scores for each country with ranking and Hofstede''s original ranks and Kendall correlations

Hofstede recognises that, across studies, scores may not be exactly the same. However the relative rankings of countries should remain reasonably consistent. To quantify the degree of agreement, Kendall's rank order correlation coefficient was used as it provides a value between –1 and +1 indicating the degree to which the two rankings agree on the orderings given to the different countries. Only one of these results suggests a strong agreement in ranking, namely, the ranks based on Individuality, and this is significant (p=0.026). Interestingly, the only other two correlations approaching significance, namely those for Masculinity and Time Orientation, are actually showing the *reverse* ordering from Hofstede!

The fact that one dimension does seem to carry over from Hofstede's study to this study is in itself good, but it is unexpected that it is the only one to do so. In order to better understand the structure of our data, we performed a factor analysis on the questionnaire data; this is discussed next.

4. FACTOR ANALYSIS OF THE VSM DATA

Hofstede's dimensions are essentially factors derived from the VSM questionnaire data. If these factors are robust, a similar analysis of our data should result in factors that closely resemble the original VSM dimensions. Hofstede [9] points out that ideally repeat analysis should use at least 10 countries - whereas we have only nine. If the factor structure is robust, however, then a smaller set of countries is likely to reveal similar factors, but of course these factors may be conflated to produce amalgams of the original VSM dimensions.

	1	2	3	4	5
IDV - q1	**0.42**	-0.02	-0.11	-0.06	**-0.53**
IDV - q2	**0.60**	0.08	-0.04	-0.11	-0.19
IDV - q4	**0.67**	-0.17	0.12	0.00	-0.07
IDV - q8	**0.51**	**0.33**	**-0.32**	-0.12	0.12
TO – q10	**0.53**	-0.24	**0.44**	-0.13	-0.05
TO – q12	**0.50**	-0.11	**0.49**	-0.03	-0.22
MAS - q15	-0.02	**0.62**	0.13	0.12	-0.18
MAS - q20	-0.01	0.15	**0.37**	-0.13	**0.66**
MAS - q5	**0.60**	0.15	-0.15	0.14	0.15
MAS - q7	**0.57**	0.07	-0.22	-0.17	0.16
PDI - q14	0.03	**0.50**	0.28	**-0.33**	-0.13
PDI - q17	0.24	0.02	**0.35**	**0.43**	0.21
PDI - q3	**0.55**	0.02	-0.28	0.24	0.16
PDI - q6	**0.65**	0.17	-0.14	0.05	0.21
UAI - q13	-0.12	**0.62**	0.23	-0.29	-0.08
UAI - q16	-0.20	**0.47**	-0.07	**0.34**	0.02
UAI - q18	0.02	0.18	0.11	**0.68**	-0.22
UAI - q19	0.11	-0.10	**0.54**	0.13	0.05

TABLE 2: VSM Principal Components Analysis for all countries

Principal Component Analysis was run on the questionnaire data from all nine countries with direct oblimin rotation in order to reveal the underlying simple factor structure [11]. As the samples were large (well over the suggested minimum 100 respondents suggested) the cut-off for significant loading of 0.3 was used [11]. All those variables

that loaded above 0.3 or below –0.3 are shown in bold in Table 2. The left hand column lists the abbreviations for each index (IDV, Individualism; TO, Time Orientation; MAS, Masculinity; PDI, Power Distance; UAI, Uncertainty Avoidance) and the number of the question as it appears on the VSM94 (e.g., q1). The numbers in bold indicate the major constituent questions of a given factor. The expected picture from this process would be that there is an initial omnibus factor followed by separate factors that reflect the VSM dimensions or that are conflations of two or more of these dimensions (given that we have fewer than 10 countries). Whilst there does seem some sort of omnibus first factor, there is very little in the rest of the table to suggest any marked similarity between the VSM dimensions and the factor structure of our data. Oddly Uncertainty Avoidance does not load on the first factor, and two questions from each of Masculinity and Power are also missing from the first factor. Additionally, none of the other factors strongly match with *any* of the VSM dimensions. The only possible exception to this is Time Orientation (in factor 3) but this does not match with Hofstede, as his loading coefficients reflect negative correlation between questions 10 and 12 and the Time Orientation factor, whereas factor 3 indicates positive correlation between these questions and that factor. There is a similar problem with the Individualism dimension, having a mix of positive and negative loadings in Hofstede's [9] equations but only having positive loadings in factor 1 for our data.

It is possible that the lack of expected loadings is due to some error in the way the VSM was administered. The VSM questionnaire was translated into several other languages: Arabic, Czech, Greek and Malay, and it is possible that the translated versions were not working "as they should" so two more factor analyses were conducted using data from just those countries sampled with the English version of the questionnaire. The first analysis included all the countries sampled in English: India, New Zealand, South Africa, UK and the USA. This produced results that were no better than those seen in Table 2.

The second analysis excluded India and South Africa. In both these countries English is used in teaching and business but is not always used in the home. English can thus often be a "second" language for many. Perhaps the questionnaire had not worked as expected because of some complication in interpreting the language used in the questionnaire in India and South Africa? This seemed unlikely, as both India and South Africa were sampled by Hofstede, but a further analysis was run to rule this possibility out. The results, as can be seen in Table 3, still do not show any of the VSM dimensions emerging as strong features in any of the factors. This suggests that whatever the problem with our use of VSM, it is not solely due to the translation process.

	1	2	3	4	5	6
IDV - q1	**0.41**	0.03	-0.28	-0.14	**-0.49**	0.11
IDV - q2	**0.49**	0.23	0.07	-0.18	-0.12	0.19
IDV - q4	**0.65**	-0.18	0.03	-0.10	0.04	0.13
IDV - q8	**0.43**	**0.50**	-0.20	-0.17	**0.31**	0.18
TO - q10	**0.43**	**-0.49**	**0.36**	-0.01	-0.01	0.05
TO - q12	**0.45**	-0.22	**0.34**	0.03	-0.17	**0.45**
MAS - q15	0.05	**0.54**	0.21	0.17	-0.23	0.09
MAS - q20	-0.12	0.00	**0.52**	-0.22	**0.45**	0.09
MAS - q5	**0.63**	0.06	0.07	0.19	-0.08	**-0.47**
MAS - q7	**0.59**	0.13	-0.12	**-0.31**	**0.37**	0.10
PDI - q14	0.03	**0.31**	**0.46**	-0.14	**-0.44**	0.03
PDI - q17	-0.01	0.10	**0.38**	**0.37**	**0.39**	0.18
PDI - q3	**0.55**	0.04	-0.11	**0.44**	0.04	-0.27
PDI - q6	**0.63**	0.07	-0.12	0.21	0.16	-0.12
UAI - q13	0.05	**0.44**	**0.44**	**-0.34**	-0.13	**-0.32**
UAI - q16	-0.16	**0.42**	0.18	0.16	0.17	-0.21
UAI - q18	-0.03	0.12	0.16	**0.67**	-0.14	0.30
UAI - q19	0.15	**-0.49**	**0.41**	-0.04	-0.04	**-0.34**

TABLE 3: VSM Principal Components Analysis for primary English speaking countries

5. OTHER POSSIBLE CAUSES FOR UNEXPECTED VSM LOADINGS

It is difficult to say why the VSM dimensions do not emerge in our dataset. It seems safe to rule out translation as a problem since the English-language only samples do not show any better factor loadings than the mixed language set. It may be possible that this is due to some other aspect of the data set, such as age. Age does influence some VSM dimensions, UAI and MAS for example [2]. Education level could also be contributing to some of the data peculiarity. Hofstede himself [9] cites work done to correct for education level for the various dimension scores, but he does not mention this as a problem for factoring the raw data. Hence neither of these issues can be completely ruled out. Also, Hofstede [9] suggests that the VSM94, which was used for this research, had not been employed enough to prove its validity without a doubt. Likewise, he suggests (in [9]) that at least 10 countries be used for a truly reliable cross-country test, whereas the present research only has 9. However, as discussed earlier, some semblance of the VSM dimensions would have been expected to emerge in our data set. Possibly the most

significant issue trying to replicate the factors by which Hofstede originated his dimensions is that we are not using the original questionnaires Hofstede used, but the version he suggests, the VSM94.

5 CONCLUSIONS

This study conducted a straightforward administration of the VSM questionnaire, yet it was unable to replicate any of Hofstede's original dimensional distinctions - with the possible exception of individuality. Factor analysis suggested that the VSM dimensions had very little explanatory power in explaining the structure of our large dataset. Our results call into question what validity the VSM model has both in itself and as a tool for understanding the design of user interfaces for different cultures, although it remains a useful shared language for discussion. The educational background and age group of the participants are possible issues in our study, but we are confident that the translation of the questionnaire is not an issue. And even if it were, the question still is *when* can VSM be used as a reliable indicator of cultural differences? At best, we can say that the VSM was measuring something — but what it was measuring is as yet unclear.

Obviously more research is needed in this area to determine what the cultural factors are that are relevant to good interaction design, and which would support HCI research in general. It may not be enough to observe the difference in interactions and interfaces from one culture to the next and to explain these observation in the light of a cultural model. There is much to be done to understand exactly what help cultural models can be to HCI.

ACKNOWLEDGEMENTS

Lidia Oshlyansky did the survey and analysis and is the main author of this paper. She has an Overseas Research Student Award and a PhD scholarship from Swansea University. Harold Thimbleby has a Royal Society-Research Merit Award. These sources of funding are gratefully acknowledged in the support of this research. Also, a great amount of gratitude goes to all the people who helped translate the questionnaires and gather the data in the countries sampled, unfortunately too numerous to list here; thank you!

REFERENCES.

[1] Hofstede, G. (1991). *Culture and Organizations: Software of the Mind*. McGraw-Hill,New York.

[2] Hofstede, G & Hofstede, G. J. (2005). *Culture and Organizations: Software of the Mind*. McGraw-Hill, New York.

[3] Marcus, A. & Gould E. (2000). Crosscurrents: Cultural Dimensions and Global Web User-Interface Design. *Interactions*, **7**, 33–46.

[4] Ford, G. & Gelderblom, H. (2003). Effects of Culture on Performance Achieved through the use of Human Computer Interaction. *Proceedings of the 2003 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologies on Enablement through Technology* (SAICSIT) Stellenbosch, South Africa, 4-6 October, 218–230.

[5] Gould, E. W., Zakaria, N., & Yusof, S. A. M. (2000) Applying culture to website design: a comparison of Malaysian and US websites. *Proceedings of the 18th Annual ACM International Conference on Computer documentation: technology & teamwork*. Cambridge, Massachusetts, 24–27 September, 161–171.

[6] Ratner, C. & Hui, L. (2003). Theoretical and Methodological Problems in Cross-cultural Psychology. *Journal of the Theory of Social Behavior*, **33** 67–94.

[7] Baskerville, R. (2003). Hofstede never studied culture. *Accounting, Organizations and Society*. **28**, 1–14.

[8] McSweeney, B. (2002). Hofstede's Model of National Cultural Differences and their Consequences: A Triumph of Faith - A Failure of Analysis. *Human Relations*. **55**, 89–118.

[9] Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and organizations across nations* (2nd edition). Thousand Oaks, California: Sage Publications, Inc.

[10] Neuman, W. L. (2000) *Social Research methods: qualitative and quantitave approaches* 4th edition. Allyn and Bacon, Boston.

[11] Kline, p. (2002). *An easy Guide to Factor Analysis*. Routledge, New York.