Are Hofstede's dimensions valid? A test for measurement invariance of uncertainty avoidance
Schmitz, Lena; Weber, Wiebke

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Are Hofstede's dimensions valid? A test for measurement invariance of Uncertainty Avoidance[1]

Sind die Hofsted'schen Kulturdimensionen valide? Ein Messäquivalenztest der Hofsted'schen Unsicherheitsvermeidungsdimension

Lena Schmitz

Dr., Wissenschaftliche Mitarbeiterin am f-bb (Forschungsinstitut Betriebliche Bildung) in Berlin. Promotion in International Cultural and Business Studies an der Universität Passau (2013)

Wiebke Weber

Dr., Leiterin der Forschungsgruppe European Social Survey (ESS) am Research and Expertise Centre for Survey Methodology (RECSM) und Mitglied des Core Scientific Teams des ESS. Promotion in Politikwissenschaft und Umfragemethodik an der Universitat Pompeu Fabra, Barcelona (2013)

Abstract (English)

On the base of a survey among IBM employees in the 1970s, Geerd Hofstede developed a model which aims at comparing cultures of countries by means of originally four, by now six dimensions. This model has evoked extreme and opposed reactions: Many researchers use it as a paradigm for cross-national comparisons, while others criticise it harshly. One basic point of criticism refers to the validity of the dimensions. The present study gives an empirical contribution to the mostly theoretical discussion and conducts tests for the validity of one Hofstedean dimension, namely Uncertainty Avoidance. Employing original data from France and Germany (2011), this dimension does not prove to be a valid construct.

Keywords: Hofstede's cultural dimensions model, Uncertainty Avoidance dimension, measurement invariance, multigroup confirmatory factor analysis, validity

Abstract (Deutsch)

Auf Basis einer Umfrage unter IBM Mitarbeitern in den 1970er Jahren konzipierte Geerd Hofstede ein Modell, das darauf abzielt, nationale Kulturen anhand von vier, mittlerweile sechs Kulturdimensionen zu vergleichen. Dieses Modell hat extreme Reaktionen in beide Richtungen hervorgerufen: Während einige Forscher ihm einen paradigmatischen Stellenwert zuschreiben, wird es von anderen scharf kritisiert. Einer der hauptsächlichen Kritikpunkte betrifft die Validität der Kulturdimensionen. Die vorliegende Studie liefert einen empirischen Beitrag zur vorwiegend theoretischen Diskussion und führt Validitätstests an einer der Dimensionen, namentlich Unsicherheitsvermeidung, durch. Unter der Anwendung von Daten aus Frankreich und Deutschland (2011) bestätigt sich diese Dimension nicht als valides Konstrukt.

Schlagwörter: Hofstedes Kulturdimensionenmodell, Unsicherheitsvermeidung, Messäquivalenz, konfirmatorische Faktorenanalyse, multipler Gruppenvergleich, Validität

1. Introduction

In the 1960s, a new sub discipline of general psychology became institutionalized, called cross-cultural psychology (Jahoda and Krewer 1997: 3, 24). Until today, researchers of this sub discipline have been following the aim of "comparing data from several cultures" (Triandis 1997:ix) in order to detect intercultural differences, usually by means of standardized questionnaires (Smith and Schwartz 1997:81). One of the most famous examples is the work of Geert Hofstede (Barmeyer and Genkova 2010:122, Harrison / McKinnon 1999:485). His comparison of *national cultures*[2] is based on cultural dimensions, which serve as standards of comparison. For a high number of countries (Hofstede 2009:xix), Hofstede calculated scores in different dimensions, i. e. numeric values, which allow establishing international rankings and country clusters (Hofstede 2009:150). His approach has been evoking extreme reactions, in a positive as well as in a critical sense. One body of research uses it as a framework for a high number of cross-cultural research projects (Blodgett et al. 2008:762, Barmeyer 2010:87). The list of areas in which the model is employed, contains a great variety[3], ranging from information technology (Myers / Tan 2002:25), management controlling (Harrison / McKinnon 1999:485), innovation, leadership styles (Jones 2007:6), over intercultural relations, decision-making, selection, training, job design, motivation and human resource management to marketing (Søndergaard 1994:453f.) and market research. Hofstede himself realizes: "Since the later 1980s the idea of dimensions of national cultures has become part of [...] 'normal science'" (Hofstede 2002:2). Accordingly, he calls his book *Culture's Consequences. Comparing Values, Behaviors, Institutions, and Organizations Across Nations* (1. edition 1980, 2. edition 2009, cop. 2001) a classic (Hofstede 2009:xvii). Thus, his dimension model has been cited, reviewed, replicated (Søndergaard 1994:447, Nakata 2009:3, Taras / Steel 2009:41) and used as a paradigm (Taras / Steel 2009:53). However, so

argue those challenging the model, the paradigmatic use has not been reflected in a sufficiently critical way. The economist Brendan McSweeney, one of Hofstede's most famous critics, observes an "on-going unquestioning acceptance of Hofstede's national culture research by his evangelized entourage" (McSweeney 2002b:112).

The present study retraces the discrepancy of the extreme reactions to Hofstede's model. We suspect that it is the debate about the validity of the cultural dimensions which essentially contributes to the discrepancy. Can the Hofstedean dimensions really serve as international standards of comparison? On which methodic foundation are they based? To contribute to this debate, the remainder of the present article proceeds as follows: First of all, the origins of the dimensions will be retraced. In chronological order, the different steps will be outlined, on the basis of which the dimensions were composed and filled with content. Thereafter, we will present follow-up studies to the Model of Cultural Dimensions. Secondly, Hofstede's statement about the dimensions' validity will be recapitulated. His statement will then be confronted with the criticisms it evoked. This confrontation leads to the research questions of the present study, if Hofstede's cultural dimensions can be considered valid constructs or not. We test this hypothesis with regard to one of the dimensions, the Uncertainty Avoidance dimension. Therefore, we replicate Hofstede's original questionnaire in a survey interrogating German and French public school teachers and factory workers, and test for measurement equivalence. In the final section we present the results and comment on their implications.

2. Hofstede's cultural dimensions theory

The starting point for the development of Hofstede's cultural dimensions theory is best summarized by the following quotation of Geert Hofstede and his son, Gert Jan Hofstede: "In the late 1960s Geert accidentally became

interested in national cultural differences – and got access to rich data for studying them" (Hofstede / Hofstede 2005:ix). The coincidence was Hofstede's involvement in developing and conducting a survey for IBM, a U.S. American multinational technology and consulting company, that aimed at studying the "job attitudes" (Hofstede et al. 1976:4) and respectively "employee values" (Hofstede 2009:41) of its employees around the world. Thus, when the IBM-questionnaire was designed, the idea of cultural dimensions was not present yet (McSweeney 2002b:95, Hofstede 2009:45). This is why Hofstedes acknowledges with regard to one of his dimensions, namely the Uncertainty Avoidance dimension that

> "It is possible that other and perhaps better survey indicators of national levels of uncertainty avoidance can be developed, but I had to use the data available in the IBM archives, and uncertainty avoidance was not a familiar concept to us when we composed the IBM questionnaire in 1967" (Hofstede 2009:148).

Between 1967 and 1973 the survey was conducted in two rounds and was completed by 160,000 employees from 72 countries in 20 languages (Hofstede 2009: 41). In the course of the analysis of the obtained data, Hofstede and his colleagues found that in some cases the different departments within one country showed stronger variations than equivalent departments of different countries (Hofstede et al. 1976:20). Regardless of this finding, Hofstede decided to focus on differences between countries, when he deduced his cultural dimensions from this study. He applied an explorative factor analysis (Hofstede 2009:31) in order to detect the underlying relationships between the given answers. This way, he found three factors. One of them he split and then created four dimensions[4]. The obvious lack of statistical independence is one of the points of criticism regarding the dimensions (Behrens 2007:71).

Hofstede chose to create exactly these four dimensions[5] to be able to draw connections to the so-called "standard analytic issues" by the sociologists Alex Inkeles and Daniel Levinson (Hofstede 2009:31). Each *standard analytic issue* represents one universal problem, that any kind of human society has to cope with (Hofstede 1998:10, Hofstede 2009:xix) and each dimension is meant to provide a bipolar (Bond / Smith 1993:41, Briley 2009:183f., Gröschke 2007:41, McSweeney 2002b:105) solution spectrum for one issue. People of different countries, so the Hofstedean argument, choose differently from the solution spectrums and, therefore, on average they can be positioned at different points between the two opposite endpoint of each dimension[6]. Hofstede names his dimensions as follows: 1) Power Distance, 2) Individualism / Collectivism, 3) Masculinity / Femininity and 4) Uncertainty Avoidance. 1) Power Distance composes the spectrum of solutions to the universal problem of human inequality and describes the extent to which members of a country accept and expect that power is distributed unequally. The higher the score of a country in the Power Distance index, the higher is the acceptance of unequal power distribution. 2) Individualism / Collectivism is related to the universal task of individuals to integrate themselves into primary groups. Whereas the members of individualist countries prefer rather loosely-knit social frameworks, the members of collectivist countries rather rely on tightly-knit ones. 3) The Masculinity / Femininity index tells if typically assumed male or female character traits predominate in the respective country. In masculine cultures, people live out the contrast between male and female traits more than in feminine cultures; this means that men show themselves especially competitive and achievement-oriented whereas women are particularly tender and socially oriented. 4) The Uncertainty Avoidance score reflects how the society of a country copes with the uncertainty of the future. The higher the score, the stronger the members of the respective country intend to avoid ambiguity (Hofstede 2009:xix f.,29).

After detecting and naming them, in a following step Hofstede intends to prove the validity of the dimensions.

For this purpose, he points out correlations between the national indices and external data (Hofstede 2009:41). As external data, he includes the findings of empirical studies, anecdotes and personal experiences (Hofstede 2009:27) as well as socio-economic indices (Hofstede 2009:68f.). The number of data considered is high: "The count of significant and independent correlations has grown to more than 400" (Hofstede 2009:4). The approach of correlating dimension indices with external data can best be demonstrated by an example. Regarding the Uncertainty Avoidance dimension, Hofstede quotes a study which compared the speed limits in fourteen countries. He discovers that those countries with stricter speed limits tend to be the ones which in the IBM study achieved a relatively low Uncertainty Avoidance score. He interprets this correlation as follows: "The emotionality in high-UAI cultures produces a sense of stress, of urgency, which in turn leads the people in those cultures to want to drive faster" (Hofstede 2009:174). This example offers to illustrate the arguments of Hofstede's critics, who call the variety of connections, which Hofstede draws to external data, boundless (Baumgartel / Thomas 1982:192) and nearly intuitive (Behrens 2007:150). They state that he compares studies which are so differently conceptualized that their findings are hardly comparable (Behrens 2007:15). Furthermore, they suspect that he only mentions those studies which fit well in his picture (Behrens 2007:56f., Early 2009:31f., McSweeney 2002a:1366). Finally, the interpretation of found correlations seems doubtful (McSweeney 2002a:1366ff.). Regarding the example above, he interprets, that people tend to drive faster in order to avoid uncertainty. An alternative equally plausible interpretation is that strong Uncertainty Avoidance finds its manifestation in strict speed limits, because the limits decrease the risk of sustaining an accident. Thus, the direction of causality between the average level of Uncertainty Avoidance and speed limits in a given country is not clear. In fact, one may wonder if there is a causal relationship at all

Area of Live	Low score on the Uncertainty Avoidance dimension	High score on the Uncertainty Avoidance dimension
Expression of Emotions	Expressions have to be controlled.	Expression of emotions normal
Subjective Well-Being (Happiness)	More subjective well-being	Less subjective well-being
Trust	Most people can be trusted	One can't be careful enough with other people, not even with family
In the Family	Lenient rules on what is dirty and taboo	Tight rules on what is dirty and taboo
At School	Dialect speech positively valued	Dialect speech negatively valued
In Motivation	Hope of success	Fear of failure
In the Work Situation	Top managers involved in strategy	Top managers involved in operations
In Consumer Behavior	Main car bought second-hand	Main car bought new
In Political Systems	Strong interest in politics	Weak interest in politics
In Legislation	Citizens positive towards legal system	Citizens negative towards legal system
Nationalism	Proud of own nation, willing to fight for it	Not proud of own nation, unwilling to fight for it
Xenophobia	Immigrants tolerated	Immigrants should be sent back

Table 1: Extract of Hofstede's typologies regarding the Uncertainty Avoidance dimension. Source: Hofstede 2009:160f., 169f., 180f.

because "correlation is not causation" (Holland 1986:945).

On the basis of the correlations, Hofstede assigns content to each dimension.

The content is presented in, as he labels them, typologies. Typologies describe those countries with a high score in the respective dimension in contrast to the ones with a low one. For both cases, typical respectively common traits and behavior are indicated, usually worded as opposites. The traits and behavior refer to all imaginable areas of life, such as work and family life, child and school education, politics, beliefs and philosophy. Table 1 illustrates a brief extract of Hofstede's typologies with reference to the Uncertainty Avoidance dimension, indicating how national cultures are in the respective areas of life if they score low on the Uncertainty Avoidance dimension compared to those that score high.

In this extract of typologies, it becomes obvious, that the Uncertainty Avoidance construct is a quite multifaceted one. If the great variety of its facets is contradictory to the construct validity, remains to be seen. In any case, Hofstede's typologies evoke criticism. One critical argument refers to their wording. One typology for instance states that in countries with a low Uncertainty Avoidance score "People feel happier" (Hofstede / Hofstede 2005:181). Unfortunately, the key term *happy* is quite wide. Moreover, some typologies lack transferability to reality. For instance, the typology regarding work life says that members of countries with a high score in the Uncertainty Avoidance index avoid taking decisions and prefer extra structured work routines. In contrast, the members of countries with a low score work better with less structured routines. Consequently, adapted at country level, this typology says that, for example, Greeks need more structured routines than Swedes (Jones 2007:6). But regardless of nationality, so we counter, specific work tasks require specific levels of routine.

The typologies, so the Linguist Leila Behrens sums up, the critical arguments, are untenable and therefore, frivolous (Behrens 2007:2, 6, 84). Still, so shall be underlined, they constitute the basic output of the Hofstedean model. This is because Hofstede assumes that the members of one country share, as they labels it, one national culture: "Culture represents the cultural mental programming that the nationals tend to have in common" (Hofstede 1980:43). He has been aiming at discovering and describing those national cultures and by creating typologies he finally means to reach this aim. Here, he transfers his findings which originally were based on a survey with specific samples, to the respective countries as a whole describing the *national culture* of these countries.[7] Consequently, it is the typologies which he recommends to be taught in intercultural textbooks and trainings (Hofstede 2002:2, Hofstede 2009:28).

Indeed, the Hofstedean model has reached an outstanding level of significance not only within the discipline of cultural sciences but also far beyond. Significance means, that, on the one hand, the model has frequently been discussed, in a critical as well as in a supportive sense, and that, on the other hand, a high number of follow-up studies have been conducted. Already in the first edition of *Culture's Consequences* (1980), Hofstede suggests continued research in the disciplines of anthropology, sociology, psychology, business, politics, law and medicine. Within these disciplines he advises to elaborate action guidelines based on his model in order to enforce its pragmatic applicability. For instance, the study conducted by Harrison et al. (1994) draws on three Hofstedean dimensions in order to explain and predict differences in organizational and management planning and control practices in Anglo-American and East Asian nations. Moreover, he recommends calculating the dimension scores of countries which were not part of the IBM study, referring in particular to socialist and development countries (Hofstede 2009:461f.). His call has been heard; the last decades brought up a great quantity of discussions and follow-up studies to his model (Søndergaard 1994:447, Hofstede 2009:66). Depending on the discipline, some dimensions have been spotlighted more than others. For instance, Individualism/ Collectivism and Masculinity/ Femininity have often been applied

within the area of psychology (Merritt 2000), whereas Power Distance and Uncertainty Avoidance were rather focused on by sociologists and business scientists (Hoppe 1990, Hastings / Hastings 1981). Later on, Hofstede revises his specification of disciplines that should apply his model: "Reviews and criticisms are most interesting when they come from unexpected areas" (Hofstede 2009:463). As unexpected areas, he mentions, among others, information technology, archive management, and nuclear power regulation. Indeed, the variety of disciplines that apply the model is great. But, so Hofstede asserts, the follow-up studies also show a great variety of quality (Hofstede 2009:66). Therefore, he dedicates a whole chapter of *Culture's Consequences*, entitled "Using Culture Dimension Scores in Theory and Research" to his requirements for continued research (Hofstede 2009:461ff.). Here, he warns of several pitfalls. Researchers should, for instance, not misunderstand the IBM questionnaire as a personality test for individuals. Furthermore, Hofstede criticises an ethnocentristic approach of those researchers who leave out the Masculinity/Femininity dimension because they consider it politically incorrect.

Now we will further analyze the different kinds of follow-up studies. It is difficult to give a representative overview covering the past 35 years, as many follow-up studies are master thesises and dissertations which either have not been translated into English or which have not been published at all (Søndergaard 1994:450). Therefore, we will concentrate on, first, the works analyzed by the Danish business scientist Mikael Søndergaard, and, second, studies mentioned by Hofstede in his texts and third some additional ones. In 1994, Søndergaard analyzed approximately 550 applications of the Hofstedean model. He had access to Hofstede's private library which allowed him to include unpublished works that had been sent to Hofstede. Søndergaard divides the applications into four categories: Citations, reviews, empirical replications and applications as a paradigm

(Søndergaard 1994:447). Hofstede slightly modifies this categorization. He divides the applications of his model into the following categories: 1) reviews and criticism, 2) extensions to new nations and regions, 3) replication studies and 4) paradigmatic uses (Hofstede 2009:461). Some examples of 1) Reviews and criticism have already been mentioned. As far as 2) extensions to new nations and regions is concerned, two examples shall be mentioned: first the master thesis by Nanhekhan (1990) and second the study by Nasierowski and Mikula (1998). The former applied the IBM questionnaire in Surinam and the latter in Poland. 3) The replications are the most frequent category of follow-up studies. "Replications are studies hat administer questions used in the IBM research to new samples from two or more of the same countries" (Hofstede 2009:463). Hofstede appreciates those kinds of studies, as he uses them to confirm the validity of his dimension constructs. Within the validation of his constructs by interpreting the correlations between his and external data, he dedicates one subchapter per dimension to "[s]traight replications of the IBM survey" (Hofstede 2009:91f., 154f., 219f., 295f.). Within straight replications, the IBM questionnaire is either applied in his original version, or the formulation of certain items is adapted to the context of the sample (Søndergaard 1994:448, Hofstede 2009:67). In any case, most replication studies follow the aim of testing Hofstede's international rankings and the majority of them confirm them (Søndergaard 1994:451). Those not confirming are usually blamed for not having adapted the items adequately to the context of the samples (Søndergaard 1994:452). As far as the samples are concerned, most replication studies do not include more than one sample per nation. They aim at detecting international differences rather than differences between subcultures. Only very few studies compare units that are smaller than nations. One study entitled "Exploring subcultural differences in Hofstede: The case of the Chinese" (Huo / Randall 1991:160) compares

samples from Taiwan, Beijing, Hong Kong and Wuhan, that means four regional subcultures facing different political situations. This kind of study though stays an exception.

> *"The disadvantages of replication and extension studies is that they are caught in the straitjacket of my model and therefore unlikely to make basic new contributions" (Hofstede 2009:465).*

Here, Hofstede leads over to the fourth category of applications: the paradigmatic uses. Hereby, he means two kinds of follow-up studies. First, he refers to those which deal with the basic structure of his model. They usually either aim at proposing an additional dimension or they might find out that one or more dimensions are not applicable in a certain region (Hofstede 2009:465). An example with much impact was the study conducted by the so-called Chinese Culture Connection (1987) which detected that the Uncertainty Avoidance was not applicable in Asia and, therefore, added the fifth dimension, Long- Versus Short-Term Orientation. Besides the studies dealing with the number of dimensions, Hofstede counts those to the fourth category of paradigmatic uses, which take the dimensions as given and base their interpretation of findings upon them.

> *"In these cases Hofstede's concepts were used as a paradigm; as a set of assumptions taken for granted. Hofstede's framework was applied in a speculative manner without any test or research based on the concepts" (Søndergaard 1994:448).*

One example for a paradigmatic use of this kind is the study by Chow, Shields and Chan (1991). They aim at investigating the influence of national culture to the high production activity of Asians and therefore compare a sample from Singapore to one from the United States. Hereby, they rely on Hofstede's assumption that they are comparing a nation with a low Individualism score to one with a high one (Chow et al. 1991:215). Some follow-up studies also combine a replication with paradigmatic use. For instance, the study by Westwood and Everett (1987) applied the complete IBM questionnaire to 170 MBA students from Malaysia, Singapore and Hong Kong (Westwood / Everett 1987:187). On the one hand, they checked, if their country scores matched those calculated by Hofstede, and, on the other hand, they conducted a factor analysis to test, if their factors were consistent with Hofstede's dimensions – which they were not (Westwood / Everett 1987:200). All in all, so Søndergaard outlines, the replication studies are the most popular kind of follow-up studies (Søndergaard 1994:450ff.). Similar to most paradigmatic uses, they rely on the Hofstedean dimensions without questioning their validity.

Also Hofstede considers his cultural dimensions as valid constructs. As outlined before, he proves this by correlating the scores countries achieve on his dimensions with external data and the interpretation of these correlations (Hofstede 2002:4, Hofstede 2009:41). Accordingly, he states a wide applicability of his dimensions (Hofstede 1998:10, Nakata 2009:4). In fact, he claims the applicability in two respects: Firstly, the dimensions are meant to be applicable in a remarkably high number of countries. However, he acknowledges that they may not be applicable in any kind of region but leave this to further research (Hofstede 2009:461). Secondly, the dimensions are expected to be applicable in all subsamples of these countries. This is because, as it has been mentioned before, according to Hofstede, national cultures are equally shared by all national members. Consequently, for a replication study basically any kind of subsamples can be chosen, not only IBM employees. Hofstede encourages replication studies (Hofstede 2009:461ff.). What he sets as a requirement, is that the nations to be compared are represented by matched samples. "We can compare Spanish nurses with Swedish nurses, or Spanish policemen with Swedish policemen" (Hofstede 2009:23). Hofstede underlines the need for equivalence among the samples; they should only differ in nationality. In case that the equivalence cannot be secured, he advises to build matched samples of more than one occupational group. "If the differences we find in one sample set

are confirmed by those found by others in other matched samples, our matching was adequate" (Hofstede 2009:23).

Whereas Hofstede is convinced of the validity of his dimensions, his critics remain skeptical: "The use of Hofstede's dimensions [...] raises more problems than it solves" (Baskerville 2003:10). What they find problematic is the dimensions' face validity, the wording of the items in the IBM questionnaire and the specificity of the IBM sample. The face validity has been tested in the course of a study conducted in 2008. Here, two groups of altogether 157 students with experience in behavioral science were asked to match the items of the IBM questionnaire to the dimensions. For instance, the items belonging to the Uncertainty Avoidance dimension were correctly matched by only 30,4% of the one group and by only 26% of the other one. This means that only a third respectively a fourth of the respondents considered the three items as reflectors of the Uncertainty Avoidance dimension as elaborated by Hofstede. This rate is evaluated as problematically low (Blodgett et al. 2008:762). These kinds of findings or simply the reconsideration of the wording of the items, as we will do in the following section, leads scholars like Behrens (2007:96) to conclude that the Uncertainty Avoidance turns out as fiction.

A possible lack of validity could be due to the wording of the items. The items which the Uncertainty Avoidance dimension is based on are presented Table 2. Using the means of these items, Hofstede created a formula[8] to calculate the Uncertainty Avoidance indices for 50 nations and three regions, namely West Africa, East Africa and several Arab countries. In the international ranking, Greece has the highest score of 112 and Singapore has the lowest of 8. The mean is 65 and the standard deviation is 24 (Hofstede 2009:151).

From these items, a set of problems evolves. To begin with, Item A37 is double-barreled because it contains two different states of feelings: "nervous" and "tense". Therefore, it does not meet the claim of unidimensionality (Raab-

Item	Position of item within the battery	Title of battery	Request for answer	Response options
A37	18	MORE ABOUT YOUR JOB	How often do you feel nervous or tense at work?	1. I always feel this way 2. Usually 3. Sometimes 4. Seldom 5. I never feel this way.
A43	18	MORE ABOUT THE COMPANY	How long do you think you will continue working for this company?	1. Two years at the most 2. From two to five years 3. More than five years (but I will probably leave before I retire) 4. Until I retire
B60	9	ABOUT GENERAL BELIEFS	Company rules should not be broken – even when the employee thinks it is in the company's best interests.	1. Strongly agree 2. Agree 3. Undecided 4. Disagree 5. Strongly disagree

Table 2: The original items on the basis of which Hofstede calculated the Uncertainty Avoidance indices. Source: Hofstede et al. 1976.

Steiner / Benesch 2010:51; Lienert / Raatz 1998:52). Moreover, the response options of this item do not constitute fixed references points, so they are not necessarily understood in the same way by every respondent (Saris 1988). Whereas, for instance once per week could by perceived as *sometimes* by one respondent, it could be considered *seldom* by another. In turn, the categories *always* and *never* are fixed reference points, but according to recent findings they might be perceived unrealistic by the respondents (Raab-Steiner / Benesch 2010:51). Furthermore, as far as item A43 is concerned, respondents could consider more than one response option adequate. They could, for instance, plan to continue working for the same company *two years at the most*, and at the same time, until they retire. Finally, regarding item B60, the wording is long-winded and, therefore, incomprehensible. Moreover, it contains a negation. Negations increase the complexity of the item and, therefore, decrease the capability of respondents to answer in a spontaneous and intuitive way. In this case, not only the item but

	France		Germany	
	Teachers	Workers	Teachers	Workers
Sample size	103 (22.60%)	113 (24.10%)	147 (31.34%)	103 (21.96%)
Male	49%	94%	49%	93%
Tertiary education	98%	4%	100%	5%

Table 3:   Summary of sample. Source: Authors own table.

3.   Data

In order to test whether the UAI is a valid construct beyond the scope of the IBM sample, we replicated the original items which indicate this dimension. In June 2011 we interviewed 113 workers of Villeroy&Boch, manufacturer of ceramics, in France, and 106 in Germany. In the same period of time we also interviewed 103 school teachers in France and 147 in Germany. The interviews were based on a paper-pencil-questionnaire which contained a total of 43 questions and took the respondents between ten and fifteen minutes to answer it. Workers as well as teachers were asked during their break at their workplace. Table 3 describes the samples. In accordance with Hofstede's sampling requirements, we included not only one but two kinds of matched samples which are equivalent: Both samples of teachers are secondary public school teachers and both samples of workers are employed by the same company and involved in comparable production processes. Moreover, the samples are adequately matched regarding their gender and education. Statistical tests confirm that the samples in the two countries are not significantly different regarding gender and education at the one percent level (see Table 3).

Although, as has been shown before, Hofstede encourages replication studies (Hofstede 2009:461ff.), he also warns researchers of replicating the IBM items in their original wording. "Questionnaires have to be adapted to their intended respondent population, situation, and period" (Hofstede 2009:67, Hofstede 1998:20f.). We followed this advice and modified the questions in order to make them applicable to the reality of the sample. Therefore, we adapted the word *company* to school (*Schule*) for the German teachers and to institution (*Institution*) for the French workers and teachers.

4.   Method

In order to test whether Hofstede's dimension of Uncertainty Avoidance is a valid concept and goes beyond the

also the response options contain negations, what makes it even more complex (Bortz / Döring 1995:224, Lienert / Raatz 1998:53). Furthermore, the situation drawn by this item is not concrete. To some respondents *Company rules* could mean life saving security rules and to others marginal rules like the obligation that employees should hang up their coat. Therefore, it appears doubtful that all respondents have the same situation in mind and give comparable answers (Bortz / Döring 1995:224, Lienert / Raatz 1998:53, Raab-Steiner / Benesch 2010:51).

Consequently, Hofstede's critics doubt the wide applicability of his dimensions. They suspect that the dimensions are IBM-specific constructs and, therefore, only serve as standards of comparison among IBM employees (Korman 1985:244, Søndergaard 1994:449). If Hofstede had interrogated a different level of employees (Hansen 2009:15), the employees of a different company (Janzer 2007:29, Goodstein 1981:51, Bond / Smith 1993:42) or even of a different profession (Cavusgil / Das 1997:216f.), so they conclude, dimensions of a different kind would have emerged. These considerations originate our research hypothesis, whereby we focus on the UAI: *The Uncertainty Avoidance dimension does not prove to be valid beyond the scope of the IBM sample.* If this hypothesis cannot be rejected, then there seems to be no foundation of this dimension, and, consequently, nor for its purpose to allow cross-cultural comparisons nor for the related typology.

scope of IBM study, we test for measurement invariance. Measurement invariance means that individuals' answers are not dependent on their group characteristics (Mellenbergh 1989, Meredith / Millsap 1992, Meredith 1993), and, thereby, that the concept is valid beyond specific groups. Figure 4 illustrates the measurement model of the Uncertainty Avoidance dimension.

In Figure 1 *Uncertainty Avoidance* is the unobserved latent concept which is reflected by the three indicators A37, A43 and B60. λ_i is the loading, τ_i is the intercept and e_i is the disturbance terms for the ith item. It is assumed that the disturbance terms have a mean of zero, and are uncorrelated with each other and with the latent variable. Moreover, in order to assign a scale to the latent variable the loading λ_2 is fixed to 1 and the intercept τ_2 to 0.

There are three different levels of invariances testing, in order: configural, metric, and scalar invariance. Configural invariance is achieved if the measurement model fits the data well in the different groups and all item loadings are significant (Davidov 2008:37). We will test for configural invariance in each group separately. Only if this is achieved, metric invariance can be tested. Metric invariance requires that the loadings λ_i are the same across groups and is a necessary condition for comparing relationships of the latent variable *Uncertainty Avoidance* with other variables. If metric invariance is established, in a final step scalar invariance can be tested. Scalar invariance requires that the intercepts τ_i are also equal across groups and if established allows comparing means across groups (Horn 1983, Meredith 1993, Steenkamp / Baumgartner 1998). Recalling Hofstede's approach, this means that only if this final condition is met, Hofstede's formula can be used and groups can be compared.

In order to conduct these tests we employ multigroup confirmatory factor analysis (MGCFA) (Jöreskog 1993, Billiet 2002) the maximum likelihood estimator of LISREL 8.72 (Jöreskog / Sörbom 2005). For model evaluation
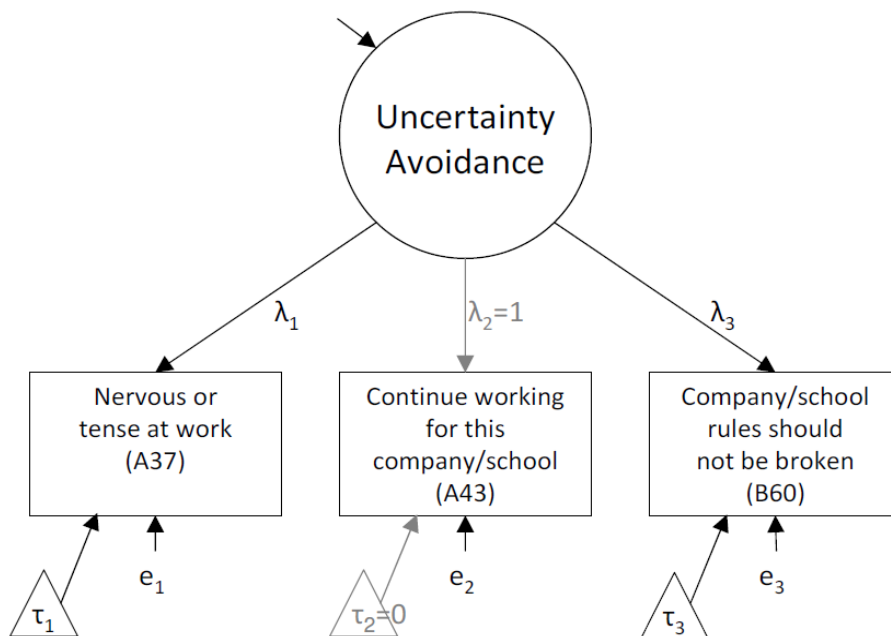

Figure 1: Measurement model of Uncertainty Avoidance. Source: Authors own figure.

and testing we rely on Jrule software (Van der Veld et al. 2008) based on the procedure developed by Saris, Satorra and van der Veld (2009). Saris et al. show that the commonly used evaluation procedures for structural equation models cannot be trusted as the test statistics and Fit indices are unequally sensitive for different misspecifications. They propose using the modification index (MI) as test statistic for detection of misspecifications (expressed as expected parameter change; EPC) in combination with the power of the MI test. This means that Jrule tests the model on the parameter level rather than the model as a whole.

5. Results

The results can be summarized rather quickly because in none of the groups configural invariance was found. In other words, in none of the groups the items load on the latent concept *Uncertainty Avoidance*. Therefore, testing for metric and scalar invariance has become redundant, as both require configural equivalence as a precedent condition. Table 4 summarizes the non-significant effects conducting factor analysis in each group separately; it presents the standardized solution which allows

		Nervous or tense at work (A37)		Continue working for this company/school (A43)		Company/school rules should not be broken (B60)	
		Unstan-dardized solution	Standardized solution	Unstan-dardized solution	Standardized solution	Unstan-dardized solution	Standardized solution
French	Teachers	.18 (1.21)	.16	1.0	1.32	-.10 (.65)	-.08
	Workers	-.51 (2.91)	-.16	1.0	.52	.28 (1.59)	.11
German	Teachers	- .87 (1.79)	-.32	1.0	.30	-.53 (.91)	-.19
	Workers	.49 (.91)	.15	1.0	.37	1.19 (2.84)	.45

Table 4: Findings of the test for configural invariance in each group. Source: Authors own table.

comparing the loadings within a group as well as the unstandardized solution which allows comparing the loading across groups. We report this for the sake of completeness, although the non-significant loadings make all comparisons of loadings redundant. The items which were supposed to reflect the Uncertainty Avoidance dimension according to Hofstede, do not act as expected.

6. Discussion

In this study, we aimed to test the validiy of the *Uncertainty Avoidance dimension* of the Hofstedean model of Cultural dimensions. We traced back the origins of this model, highlighted the importance and use of it, mentioned the criticisms to it, and tested with original data from France and Germany (2011), if the Uncertainty Avoidance dimension can be replicated and hence, proves to be a valid construct beyond the original IBM study. We find that this is not the case. In none of our four samples, the items have proven to be indicators of the construct *Uncertainty Avoidance*. This contradicts Hofstede who claims that his dimensions are applicable not only in a high number of nations, but also among all subsamples

within these nations. From our results we conclude, that criticisms suspecting that the Uncertainty Avoidance dimension is specific to the Hofstedean IBM sample are justified. Consequently, this dimension is neither to be used as a standard of cross-national comparisons, nor as the basis for general descriptions about countries as wholes.

The lack of validity we find leads to the question what it is due to. One theoretical explanation would be cultural change over time. Uncertainty Avoidance might have been a valid construct in the 1970s and might have lost relevance over the years. This might be the reason for our sample from 2011 to not replicate it. This explanation though does not convince in two respects. First, Hofstede himself excludes the idea of cultural change. "Cultures, especially national cultures, are extremely stable over time" (Hofstede 2009:34). More concretely speaking, national cultures are supposed to be stable "across many generations" (Hofstede 2009:10), at least until the year 2100 (Hofstede 2009:36). Whenever his data is blamed to be out of date, Hofstede points to the century-old roots of his dimensions (Hofstede 2009:73). Here, he mentions the second objection to the explana-

tion of cultural change over time. If the dimensions really do reflect universal problems of every human society and if country scores shall be published and discussed until today, then it is indispensable that the dimensions have not lost their validity.

In the beginning of the present study, we recapitulated in detail the foundation and methodological approach of the Hofstedean model. Thereby, we detected several reasons which could lead to the lack of validity of dimension we studied here. First of all, the aim of IBM questionnaire was not to compare cultures across countries and the idea of the Uncertainty Avoidance dimension was not present at all when the questionnaire was designed. Secondly, once the data from the IBM study was obtained, Hofstede split the three factors which he found by means of factor analysis into four dimensions. The theoretical reason was to bring the dimension in line with the *standard analytic issues* by Inkeles and Levinson (1954). However, practically this means that the dimensions are not statistically independent from each other. Thirdly, given that the aim of the questionnaire design was not to detect the Uncertainty Avoidance dimension, the wordings of the items are doubtful. Hofstede himself considers this problematic and recommends adjustments in the formulation of the items. However, he still believes that these three items are the indicators of the Uncertainty Avoidance dimension. This reasons seems to explain why it was impossible to detect the Uncertainty Avoidance dimension with our data.

All in all, our study made clear that the Uncertainty Avoidance dimension is not a valid construct and, therefore, does not serve as base for follow-up studies. The majority of these studies are replications of the original study, mainly, in order to update the original study with more recent data or to derive action guidelines that are practically applicable. In contrast to our replication, the majority of the studies does not question the validity of the constructs and hence, take the dimensions as given. Following from our finding that the

Uncertainty Avoidance dimension has not proven valid, those studies show a lack of sustainable foundation. Construct validity should have been assured by testing for measurement invariance across groups. As a conclusion, this does not mean that a concept such as the Uncertainty Avoidance does not exist, but it means that it cannot be measured employing (variations of) the items from the IBM study.

7. Bibliography

Barmeyer, C. / Genkova, P. / Scheffer, J. (2010): *Interkulturelle Kommunikation und Kulturwissenschaft*. Passau: Stutz.

Barmeyer, C. (2010): Kulturdimensionen und Kulturstandards. In: Barmeyer, C. / Gankova, P. / Scheffer, J. (Eds.): *Interkulturelle Kommunikation und Kulturwissenschaft*. Passau: Stutz, pp. 87-117.

Baskerville-Morley, R. (2005): A research note: the unfinished business of culture. *Accounting, Organizations and Society* 30, pp. 389-391.

Baumgartel, H. / Thomas, H. (1982): Review of Geert Hofstede: `Culture's Consequences: international differences in work-related values´. *Personnel Psychology* 35(1), pp. 192-196.

Behrens, L. (2007): *Konservierung von Stereotypen mit Hilfe der Statistik*. Köln: Institut für Linguistik, Allgemeine Sprachwissenschaft.

Billiet, J. (2002): Cross-cultural equivalence with structural equation modeling. In: Harkness, J. A. / Van de Vijver, F. J. R. / Mohler, P. P.(Eds.): *Cross-Cultural Survey Methods*. Hoboken: John Wiley, pp. 247-264.

Blodgett, J. / Bakir, A. / Rose, G. (2008): A Test of the Validity of Hofstede's Cultural Framework. *Advances in Consumer Research* 35, pp. 762 f.

Bond, M. H. / Smith, P. B. (1993): *Social psychology across cultures*. London: Harvester Wheatsheaf.

Briley, D. A. (2009): Cultural Influence on Consumer Motivations: A Dynamic View. In: Cheryl, N. (Eds.): *Beyond Hofstede*. New York: Palgrave Macmillan, pp. 181-200.

Cavusgil, S. T. / Das, A. (1997): Methodology issues in cross-cultural sourcing research - a primer. *Marketing Intelligence and Planning* 15(5), pp. 213-220.

Chow, C. W. / Shields, M. D. / Chan, Y. K. (1991): The effects of management controls and national culture on manufacturing performance: an experimental investigation. *Accounting, Organizations and Society* 19 (5), pp. 209-226.

Davidov, E. (2008): A cross-country and cross-time comparison of the human values measurements with the second round of the European Social Survey. *Survey Research Methods* 2, pp. 33-46.

Early, C. P. (2009): So What Kind of Atheist Are You? In: Nakata, C. (Eds.): *Beyond Hofstede*. New York: Palgrave Macmillan, 19-39.

Goodstein, L. D. (1981): Commentary: Do American Theories Apply Abroad? *Organizational Dynamics* (Summer), pp. 49-54.

Gröschke, D. (2007): Kulturelle Unterschiede im Selbstkonzept. *Interculture Journal* 6(5), pp. 39-70.

Hansen, K. P. (2009): Die Problematik des Pauschalurteils. *Interculture Journal* 8(10), pp. 5-18.

Harrison, G. L./ McKinnon, J. L. / Panchapakesan, S. / Leung, M. (1994): The Influence of Culture on Organizational Design and Planning and Control in Australia and the United States Compared with Singapore and Hong Kong. *Journal of International Financial Management and Accounting* 5(3), pp. 242-261.

Harrison, G. L. / McKinnon, J. L. (1999): Cross-cultural research in management control systems design: a review of the current state. *Accounting, Organizations and Society* 24, pp. 483-506.

Hastings, H. E. / Hastings, P. K. (1981): *Index to interntaional public opinion 1979-1980*. Oxford: Clio.

Hofstede, G. / Kraut, A. I. / Simonetti, S. H. (1976): *The development of a core attitude survey questionnaire for international use.* Brussels: European Institute for Advanced Studies in Management.

Hofstede, G. (1980): Motivation, Leadership, and Organization: Do American Theories Apply Abroad?. *Organizational Dynamics* (Summer), pp. 42-66.

Hofstede, G. (1998): Masculinity and Femininity as a Dimension of Culture. In: Hofstede, G. (Ed.): *Masculinity and Femininity: the taboo dimension of national cultures.* Thousand Oaks: Sage Publications, pp. 3-28.

Hofstede, G. (2002): Dimensions do not exist: A reply to Brendan McSweeney. *Human Relations* 55(11), pp. 1-8.

Hofstede, G. / Hofstede, G. J. (2005): *Cultures and organizations*. New York, N.Y.: McGraw-Hill.

Hofstede, G. (2009): *Culture's Consequences. Comparing Values, Behaviors, Institutions and Organizations Across Nations.* London: Sage Publications.

Hoppe, M. H. (1990): *A comparative study of country elites: International differences in work-related values and learning and their implications for management training and development.* Dissertation: University of North Carolina at Chapel Hill.

Horn, J. L. / McArdle J. J. / Mason R. (1983): When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. *The Southern Psychologist* 1, pp. 179-209.

Huo, Y. P. / Randall, D. M. (1991): Exploring subcultural differences in hofstede's value survey: The case of the Chinese. *Asia pacific Journal of Management* 8(2), pp. 159-173.

Jahoda, G. / Krewer, B. (1997): History of Cross-Cultural and Cultural Psychology. In: Berry, J. W. / Poortinga, Y. H. / Pandey, J. (Eds.): *Handbook of Cross-Cultural Psychology*. Boston: Allyn and Bacon.

Janzer, A. (2007): Kulturwissenschaftliche Probleme internationaler Managementforschung. *Interculture Journal* 6(5), pp. 21-38.

Jones, M. L. (2007): *Hofstede – Culturally Questionable?* Speech at the Oxford Business & Economics Conference on June, 24th-26th 2007, Oxford University.

Jöreskog, K. G. / Dag S. (2005): *LISREL 8.57.* Lincolnwood, IL: Scientific Software International.

Jöreskog, K. G. (1993): Testing structural equation models. In: Bollen, K. A. / Long,

S. J. (Eds.): *Testing structural equation models*. Newbury Park: Sage Publications, pp. 294-316.

Korman, A. K. (1985): Review of `Culture's Consequences: International differences in work-related values´. *Journal of Occupational Behaviour* 6(3), pp. 243 f.

Lienert, G. A. / Raatz, U. (1998): *Testaufbau und Testanalyse*. Weinheim: Beltz, Psychologie Verl.-Union.

McSweeney, B. (2002a): The Essentials of Scholarship: A Reply to Geert Hofstede. *Human Relations* 55(11), pp. 1363-1372.

McSweeney, B. (2002b): Hofstede's model of national cultural differences and their consequences: A triumph of faith - a failure of analysis. *Human Relations* 55(1), pp. 89-118.

Mellenbergh, G. J. (1989): Item bias and item response. *International Journal of Educational Research* 13, pp. 127-143.

Meredith, W. / Millsap, R. E. (1992): On the misuse of manifest variables in the detection of measurement invariance. *Psychometrika* 57(2), pp. 289-311.

Meredith, W. (1993): Measurement invariance, factor analysis, and factorial invariance. *Psychometrika* 58, pp. 525-543.

Merritt, A. C. (2000): Culture in the cockpit: Do Hofstede's dimensions replicate? *Journal of Cross-Cultural Psychology* 31, pp. 283-01.

Myers, M. D. / Tan, F. B. (2002): Beyond Models of National Culture in Information Systems Research. *Journal of Global Information Management* 10(1), pp. 24-32.

Nakata, C. (2009): Going Beyond Hofstede: Why We Need to and How. In: Nakata, C. (Ed.): *Beyond Hofstede*. New York: Palgrave Macmillan, pp. 3-18.

Nanhekhan, R. (1990): *Cultuurverschillen gemeten: aanknopingspunten voor wederzijds begrip tussen Nederlands en Surinamers op de werkvloer*. Amsterdam: Master Thesis.

Nasierowski, W. / Mikula, B. (1998): Culture Dimensions of Polish Managers: Hofstede's Indices. *Organization Studies* 19(3), pp. 495–509.

Raab-Steiner, E. / Benesch, M. T. (2010): *Der Fragebogen*. Stuttgart: UTB.

Saris, W. E. (1988): *Variation in Response Functions: A Source of Measurement Error in Attitude Research*. Amsterdam: Sociometric Research Foundation.

Saris, W. E. / Satorra, A. / Van der Veld, W. (2009): Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal* 16(4), pp. 561-582.

Smith, P. B. / Schwartz, S. H. (1997): Values. In: Berry, J. W. / Segall, M. H. / Kagitcibasi, C. (Eds.): *Handbook of Cross-Cultural Psychology*. Boston: Allyn and Bacon.

Søndergaard, M. (1994): Research Note: Hofstede's Consequences: A Study of Reviews, Citations and Replications. *Organization Studies* 15(3), pp. 447-456.

Steenkamp, J.-B. / Baumgartner, H. (1998): Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research* 7, S. 78-90.

Taras, V. / Steel, P. (2009): Beyond Hofstede. In: Nakata, C. (Ed.): *Beyond Hofstede*. New York: Palgrave Macmillan, pp. 40-60.

Triandis, H. C. (1997): Foreword. In: Berry, J. W. / Dasen, P. R. / Saraswathi, T. S. (Eds.): *Handbook of cross-cultural psychology*. Boston: Allyn and Bacon, pp. ix f.

Van der Veld, W. / Saris, W. E. / Satorra, A. (2008): *Jrule 2.0: User manual*. Unpublished manuscript (Internal report, Radboud University Nijmegen, The Netherlands).

Westwood, R. G. / Everett, J. E. (1987): Culture's consequences: A methodology for comparative management studies in Southeast Asia? *Asia pacific Journal of Management* 4(3), pp. 187-202.

Endnotes

1. The theoretical part of this article is based on: Schmitz, L. (2013): Nationalkultur versus Berufskultur. Eine umfassende Kritik Hofstedes. Diss. Universität Passau.

2. "In most of this book, I use the word culture to refer to national culture [...]" (Hofstede 2009:1).

3. „[...] hundreds of researchers have used the Hofstedean framework to understand culture's influences on managerial, consumer, and organizational behavior" (Adir et al. 2009:146).

4. He split one of the factors into Power Distance and Individualism / Collectivism (Hofstede 2009:59).

5. By now, the number of dimensions has extended to six (Hofstede 2011:8).

6. In general, this idea refers to the concept by Clyde Kluckhohn (Hofstede 2009:9f., 28f., Baskerville-Morley 2005:389, Nakata 2009:4, Rathje 2009:35).

7. It shall be noted, that Hofstede – even though he aims to detect national cultures ("[...] I use the word *culture* to refer to national culture" (Hofstede 2009:1)) – attaches dimension scores to nations, countries as well as to regions, including Western Germany, Belgium and Arab countries (Hofstede 2009:44, 87). In order to avoid the inherent confusion, we refer here to *cultures of countries* when Hofstede mentions *national cultures*.

8. The formula is the following: UAI = 300 – 40 x (mean score A37) – (% answer 1 or 2 in A43) – 30 x (mean score B60) (Hofstede 2009:491). According to the original IBM study, the countries involved in the present study scored as follows: France reaches a score of 86, and, therefore, the tenth rank in the UA index, whereas Germany obtained a remarkable smaller score of 65, and, therefore, occupied the 29th rank.