The use of English questionnaires in cross-national research

Anne-Wil Harzing

Version April 2004

Dr. Anne-Wil Harzing   Email: anne-wil@harzing.com
University of Melbourne   Web: www.harzing.com
Department of Management
Faculty of Economics & Commerce
Parkville Campus
Melbourne, VIC 3010
Australia

The use of English questionnaires in cross-national research:
Does cultural accommodation obscure national differences?

Anne-Wil Harzing, University of Melbourne,
Department of Management, Faculty of Economics & Commerce
Parkville campus, Melbourne, Victoria 3010, Australia
Tel: + 61 3 8344 3724, Email: anne-wil@harzing.com

Country collaborators[1]

Austria: Iris Fischlmayr, Johannes Kepler University, Linz
Brazil: Maria Ester Freitas, Escola de Administração de Empresas de São Paulo
Bulgaria: Mila Lazarova, Simon Fraser University (Canada)
Chile: Leonardo Liberman Yaconi, Queensland University of Technology (Australia)
China: Ying Zhu, University of Melbourne (Australia)
Denmark: Mikael Søndergaard, University of Southern Denmark
Finland: Rebecca Marschan-Piekkari, Swedish School of Economics
France: Cordula Barzantny, Groupe ESC Toulouse Business School
Germany: Katrin Wittenberg, University of Bradford School of Management (UK)
Greece: Barbara Myloni, Athens University of Economics and Business
Hong Kong: Wai-Ming Mak, Hong Kong Polytechnic University
India: Srabani Roy Choudhury, International Management Institute
Japan: Linda Viswat, Otemon Gakuin University
Lithuania: Audra Mockaitis, Vilnius University, Laura Salciuviene, Kaunas University of Technology
Malaysia: Janaka Chee Kong Low, University of Science Malaysia
Mexico: Maria Soledad Gomez Lopez, ITESM
Netherlands: Ayse Saka, University of Groningen, Deanne Den Hartog, Erasmus University
Poland: Aleksandra Grzesiuk, West Pomeranian Business School
Portugal: Francisco B. Castro, CEMPRE, Faculdade de Economia do Porto, Universidade do Porto
Russia: Graham Roberts, Groupe ESC Lille/Lille Graduate School of Management (France)
Spain: Ignacio Martínez, University of Navarra
Sweden: Lena Zander, Ciara Karlsson, Laurence Romani, IIB, Stockholm School of Economics
Taiwan: Whitman K. Liang, Tatung University
Turkey: Ayda Uzuncarsili Soydas, Gozde Yilmaz, Marmara University
United Kingdom: Alan Feely, Aston Business School
United States: Miguel R. Olivas-Luján, ITESM (Mexico), Martha Maznevski, IMD (Switzerland)

The use of English questionnaires in cross-national research:
Does cultural accommodation obscure national differences?

ABSTRACT

Cross-national research is plagued by many obstacles. This article focuses on one of these obstacles: the fact that research in more than one country usually involves subjects with different native languages. We investigated whether the language of the questionnaire influences response patterns. More specifically we tested the cultural accommodation hypothesis: Do respondents subconsciously adjust their responses in a way that reflects the cultural values associated with the language of the questionnaire?

We tested this hypothesis with a sample of 3,419 undergraduate students in 24 countries. Half of the students in each country received an English-language questionnaire, while the other half received the same questionnaire in their native language. Three types of questions were included in the questionnaire: questions about cultural norms and values, questions about characteristics of the ideal type of jobs that students would prefer after graduation, and questions about reasons for choosing elective subjects in their studies. Comparison with a comparison group of US students showed cultural accommodation to be present in a substantial proportion of the comparisons for all three types of questions. Consequences and recommendations for cross-national research are discussed.

Keywords: international research, language, cultural accommodation

INTRODUCTION

Cross-national research is plagued by many problems (for an overview see for instance Singh, 1995, Usunier, 1998; van de Vijver & Leung, 2000). One of these problems is the fact that when doing research in more than one country the researcher usually encounters respondents with different native languages. A researcher confronted with a linguistically diverse population can translate the questionnaire into as many languages as necessary. This would be the researcher's only option if respondents are monolingual or if there is no shared second language among respondents. Brislin (1986) offers a set of recommendations for translation of research instruments. However, questionnaire translation is not an unambiguous process and might be time-consuming and expensive. Fortunately, respondents with sufficient language capabilities might be able to respond to the questionnaire in its original language (usually English). However, this leads to another question: could the language of the questionnaire influence a person's response?

There are two different conceptions of the role of language in the study of cross-national differences that might be able to offer an answer to this question: the Whorfian and the Linguistic positions (Hulin & Mayer, 1986). According to the extreme Whorfian position, individuals who speak different languages live in different worlds, rather than living in the same world with different labels for objects, events, and concepts. This position is based on the Sapir-Whorf hypothesis that sees language as a filter between an individual and his environment. Language has such a strong impact that cross-language research is virtually impossible. According to the extreme linguistic position, very high fidelity translations from a source to a target language would provide a sufficient basis for cross-language and cross-national assessments and comparisons. According to this position, languages are simply linguistic symbols for common terms and can be translated into an equivalent set of symbols, a different language, with little loss of meaning (Brislin, 1980, cited in Hulin & Mayer, 1986).

Although neither of these positions is likely to be accurate in their extreme forms, a less extreme version of the Whorfian hypothesis suggests that the language of the questionnaire might influence

people's responses to a questionnaire. Since language and culture are interrelated, this influence is especially likely when the instrument assesses cultural norms and values. Yang and Bond (1980) have termed this process cultural accommodation. Previous research has found some evidence that the language of a questionnaire can influence individuals' responses (see for instance Earle, 1969; Botha, 1970; Bond & Yang, 1982; Candell & Hulin, 1986; Schemerhorn, 1990; Ralston, Cunniff & Gustafson, 1995; Erdener, Wongtada & Fagenson, 1996; Harzing et al., 2002). This article will focus on testing the cultural accommodation hypothesis in a more controlled setting and on a larger scale than has been done so far.

The remainder of this article is structured as follows. We will first briefly discuss previous studies that have assessed the impact of language on response patterns and implicitly or explicitly tested the cultural accommodation hypothesis. We will then describe how our study complements these studies. Subsequently, we will discuss our methodology and results. A discussion section will conclude our article.

LITERATURE REVIEW

Studies on the impact of language on response patterns have focused on one of two approaches: within-participant comparisons and between-participant comparisons (Sanchez, Alonso & Spector, 2000). The within-participant approach presents the same questionnaire in two different languages to every respondent. The between-participant approach splits up the group of respondents and each respondent answers the questionnaire in only one language.

Results for within-participant comparisons are mixed. Earle (1969) and Botha (1970) found significant differences between language versions that supported cultural accommodation. Erdener, Wongtada & Fagenson (1996) found significant differences for a sub-group of their sample, but did not test for cultural accommodation. Katerberg, Smith & Hoy (1977), Tyson, Doctor & Mentis (1988) and Sanchez et al. (2000) did not find any differences between language versions. However, it is possible that respondents will make an effort to remember their earlier responses. Separating the administration of

the questionnaire in time in order to address this problem – as was done in many of these studies – allows for confounding variables to intervene. Depending on the design of the study it might also lead to a much smaller sample, since some respondents might decline participation in the second study.

The between-participant approach eliminates the potential consistency bias, but puts heavy demands on the comparability between samples. Four studies that applied this approach (Bond & Yang, 1982; Harzing et al., 2002; Ralston et al., 1995; Schemerhorn, 1990) found differences between language versions that supported the cultural accommodation hypothesis, while one study (Candell and Hulin, 1986), found only very minor differences between language versions. Ralston et al.'s (1995) study illustrates a major drawback of the between-participant approach in comparison to the within-participant approach: it is very difficult to find samples that are matched on all other characteristics apart from the language of the questionnaire (Sanchez et al., 2000). Respondents might differ in terms of demographic characteristics, their position in the company, the type of company they work for etc. Although some of these characteristics were measured in the Ralston et al. (1995) study, they were not included in the analysis.[2]

OUR STUDY'S CONTRIBUTION

We have chosen to use the between-participant approach in our study, since we feel that the "consistency" problem associated with the within-participant approach would hinder meaningful comparisons. However, we have made every effort to avoid the problems associated with the between-participant approach by eliminating self-selection and matching respondents very closely. In addition, our study will improve upon earlier studies in several ways.

First, like Harzing et al. (2002), we include both questions that relate to cultural values and questions that are more neutral. Most earlier studies focused on only one category of questions, either cultural values (Earle, 1969; Botha, 1970; Bond & Yang, 1982; Tyson et al., 1988; Ralston et al. 1995) or questions dealing with organizational issues such as job description and organizational commitment

(Katerberg et al. 1977; Candell & Hulin, 1986; Sanchez et al., 2000). Generally, studies focusing on cultural values found a response effect, while studies focusing on more neutral questions did not.

Second, our study compares English with no fewer than 18 other languages in 24 countries. It therefore includes a much wider range of countries and languages than previous studies that usually focused on a comparison between English and one other language only: Chinese (Earle, 1969; Bond & Yang, 1982; Schermerhorn, 1990; Ralston et al., 1995), Spanish (Katerberg et al., 1977; Sanchez et al., 2000), Afrikaans (Botha, 1970, Tyson et al. 1988) or French (Candell & Hulin, 1986). Harzing et al. (2002) included six languages in seven different countries. However, this study suffered from using a comparison group of UK students that was not ideally matched with the other respondents. It also included only two non-European countries and was not able to completely rule out translation inaccuracy as only two countries in the study (Austria & Germany) shared the same language.

Our current study includes North-European (Austria, Denmark, Finland, Germany, the Netherlands, Sweden), South European (France, Greece, Portugal, Spain), Central or East European (Bulgaria, Lithuania, Poland, Russia, Turkey), Latin American (Brazil, Chile, Mexico) and Asian (China, Hong Kong, India, Japan, Malaysia, Taiwan) countries, and will therefore allow us to test the cultural accommodation hypothesis with a much larger and more varied sample. Our study's main hypothesis is reproduced below. Details of our study's design can be found in the next section.

Hypothesis 1: When replying to an English-language questionnaire non-native speakers of English will culturally accommodate their answers to the values embedded in the English language and hence their responses will be closer to those of native English speakers than when they reply in their native language.

STUDY DESIGN & METHODOLOGY

INSTRUMENT

As indicated above, the aim of our study is to assess whether the language of the questionnaire has an impact on the way people respond to the questions included in this questionnaire. Because of the inter-

action between language and culture, response differences are more likely for questions that relate to cultural values than for more neutral questions. Our instrument therefore included examples of both types of questions. Since our population consisted of students, our neutral questions related to reasons for choosing electives. In addition to cultural values and elective choice questions, a third set of questions was introduced that asked students to assess the importance of various characteristics of their ideal job after graduation. These questions might be expected to show *some* language effect, since they might refer to cultural values.

In designing the instrument we followed best practices in cross-cultural research (as recently summarized by Schaffer & Riordan, 2003), such as using back-translation, avoiding colloquial language, using short and simple sentences, repeating nouns instead of using pronouns wherever possible and considering the national collaborators' perspectives in designing the questionnaire. This section first discusses the construction of scales for each of the three types of questions. Subsequently, we discuss the translation procedures that we used to ensure conceptually equivalent translations. Finally, this section discusses the standardisation procedures that were necessary to remove the systematic response bias unrelated to the language of the questionnaire.

Cultural values scales

To measure cultural values, we used a revised version of the Cultural Perspective Questionnaire (Maznevski et. al., 2002), which is based on the culture framework presented by Kluckhohn & Strodtbeck (1961). Because of constraints in terms of questionnaire length, we chose to focus on only two of the six cultural dimensions: Activity and Relationships, each with three variations. The three variations of basic modes of Activity are *doing, being* and *thinking.* The three types of naturally occurring Relationships among humans are *individualism, collectivism,* and *hierarchy.*[3] Kluckhohn & Strodtbeck clearly identified individuals as the "holders" of the preference for variations, whereas the cultural pattern was defined by the aggregation of individuals' preferences. We can therefore make hypotheses and test them at the individual level of analysis and aggregate measures to develop descriptions of cultures.

Each of the variations was measured with 7 single-sentence items and respondents were asked to record their strength of agreement with each, on a Likert scale from 1 (strongly disagree) to 5 (strongly agree).[4] In order to verify the validity of our constructs, we conducted a factor analysis (principal components, with oblique rotation) for the aggregate sample for both the Activity and Relationship dimension, imposing a 3-factor solution for each dimension. Bartlett's test of sphericity was highly significant (6213.846, p <0.000 for Activity and 7592.044, p <0.000 for Relationships) and KMO's measure of sampling adequacy was 0.77 for Activity and 0.82 for Relationships, which is considered more than satisfactory. For the Activity dimension, items for the three variations loaded consistently on three different factors, with the exception of four items (2 for Doing and 1 each for Being and Thinking) that either did not have a significant loading on any of the factors or loaded on a different factor. However, when the item loaded on the "wrong" factor, it had a lower loading on that factor than all other items that were designed to measure the construct and usually had a secondary loading on the "right" factor. Since exclusion of these four items reduced scale reliability, we decided to keep them in their respective scales. For the Relationships dimension, items for the three variations loaded consistently on three different factors, with the exception of three items (1 for each dimension). For the Hierarchy and Collectivism variation, exclusion of these items resulted in higher scale reliability, hence they were excluded from the scales.

As Table 1 shows the resulting reliability of the Activity *thinking* scale was adequate (Cronbach's alpha: 0.72), while scale reliabilities for the Activity *doing*, Relationship *hierarchy* and Relationship *individualism* scales were lower (Cronbach's alpha: 0.59, 0.62 and 0.58 respectively). Both the Activity *being* variation and the Relationship *collectivism* variations had even lower reliabilities (0.47/0.51). Results for these two variations should therefore be interpreted with caution. Although overall, scale reliabilities are not high, we consider scale reliabilities around or above 0.60 acceptable given the range of cultures represented in our sample and the difficulty of measuring complex constructs such as culture reliably. We are supported in this by Thomas & Au (2002), who report reliabilities of 0.60 and 0.57 respectively for their measures of Individualism and Collectivism and indicate these are consistent with previous re-

search. They also argue that these relatively low reliabilities are typical of cultural dimensions, that normally have a high "bandwidth". Bandwidth is the amount of information captured in the construct and has to be balanced with fidelity, the consistency of information. According to Thomas & Au (2002), the need for this delicate balance in the measurement of culture has been widely discussed and has resulted in broad acceptance in the area of cross-cultural psychology (see also Kemmelmeier et al. 2003 and Shafiro, Himelein & Best, 2003).

Insert Table 1 about here

Elective questions scales

As a representative for neutral questions, we asked students for the rationale in their selection of electives, providing a range of eight predefined reasons. Even though answers to these questions might differ across countries, we do not expect them to differ systematically between languages in one country, since these questions would not appear to be clearly related to cultural values. The eight elective questions were subjected to factor analysis (principal components, with oblique rotation) to discover any underlying constructs. The original analysis provided a 2-factor solution, in which six questions loaded highly (0.56-0.68) on the first factor, while "because I am interested in the subject" and "because it is relevant for my future career" loaded highly (0.69-0.76) on the second factor. However, the two questions related to the lecturer ("because I like the lecturer", "because the lecturer has a good reputation") showed significant positive loadings on the second factor as well. In fact in the unrotated solution both questions loaded equally high on both factors. Further, the scree-plot only started to level off after the 3rd factor. We therefore re-ran the factor analysis enforcing a three factor solution. Bartlett's test of sphericity was highly significant (4537.949, p <0.000) and KMO's measure of sampling adequacy was 0.71, which is considered more than satisfactory. In this solution, that explained 63% of the variance, the two lecturer questions were separated out as a third factor. The first factor, which included the reasons: "because my friends are choosing it", "because it is scheduled at a convenient time", "because it is

9

less work", "because I can get a high mark for it", can be concluded to represent "extrinsic" factors, while the 2nd factor (interested and career) can be seen as representing "intrinsic" reasons to choose electives.

We therefore created scales for the intrinsic, extrinsic and lecturer-related questions. As Table 1 shows the extrinsic and lecturer-related scales have reliabilities that are acceptable, given the wide range of countries represented in our sample. The intrinsic scale has a comparatively lower reliability. However, since it consistently loaded very highly on a single factor, we have decided to maintain this as a scale. However, the results for this elective scale should be interpreted with some caution.

Ideal job scales

The third set of questions asked students to assess the importance of various characteristics of their ideal job after graduation and was adapted from Sirota & Greenwood (1971) and Hofstede (1980). A total of 18 questions were included in the questionnaire. These questions were subsequently subjected to factor analysis (principal components, with oblique rotation). Bartlett's test of sphericity was highly significant (6603.058, $p < 0.000$) and KMO's measure of sampling adequacy was 0.83, which is considered more than satisfactory. The 18 questions were reduced to four relatively clear factors. As can be seen in Table 2, the first factor included characteristics that referred to a balance between work and private life and having a job that was not very demanding, but allowed good relationships with others. It was labeled *balance and relationships*. The second factor represented mainly *job intrinsic* elements and was labeled as such. Factor 3 clearly referred to monetary rewards and advancement and was therefore labeled *money & advancement*. The final factor would seem to refer to an orientation to serve and was labeled *serving*. Scale reliabilities were subsequently calculated for each of the factors and were deemed acceptable, given the wide range of countries represented in our sample (see Table 1).

Insert Table 2 about here

10

Translation

The questionnaire was designed in English. It was pilot tested in the UK in October 2000. This pilot test resulted in the replacement of some items for the cultural dimensions and the introduction of the ideal job questions. The pilot study coincided with a discussion among the first eight country collaborators about translatability of items. Several items that proved to be difficult to translate were replaced. Subsequently, bilingual country collaborators were responsible for the translation of the original English questionnaire. Translations were conducted using translation-back-translation procedures. The translator and back-translator where separate individuals who did not enter into a discussion until after they had finished their translations. Discussions between translator and back-translator usually resulted in the change of some of the translations. Where difficulties remained, a third bilingual person was consulted. The back-translated versions were verified by the project coordinator for consistency across languages, which usually resulted in further changes and discussions between translator and back-translator. For several of the European languages the project coordinator provided independent verification of the translated versions.

As a test of translation accuracy we compared scale reliabilities between the two language versions (native, English). If there had been a systematic problem with the translation, reliabilities for the native language versions would have been lower than for the English language versions. Our results show this is not the case. Reliabilities for the two language versions were virtually identical for the cultural dimension scales, except for the Relationship *individualism* scale where the reliability for the native language version was 0.02 higher. For the elective questions, reliabilities did differ for all three scales by 0.01-0.05. In all cases, however, reliability for the native language version was higher. Reliabilities for two of the four ideal job scales were virtually identical across languages, while the other two differed by 0.05-0.07. Again though, the native language versions showed a higher reliability.[5]

Although, as in any multi-country study, it is very difficult to guarantee translation accuracy with absolute certainty, we are quite confident that the resulting questionnaires are equivalent in meaning across languages. Further, any potentially remaining translation inaccuracies would be attenuated by

two factors. First, apart from two of the elective scales, scales generally consisted of 3-7 items and hence translation inaccuracies in one item would not have a major impact on overall results. Second, we are looking at overall patterns with regard to cultural accommodation across 24 countries, so that any translation inaccuracies for specific languages would not have a major impact on overall results.

Systematic response effects

In this study we are interested in the response effect caused by the respondent answering in different languages. However, previous research has demonstrated a significant country effect on respondents' tendency to use different parts of the scale (Leung & Bond, 1989; Mullen, 1995; Singh, 1995). Since this would impact on our between-country comparisons, it is important to assess whether these response effects are present in our sample. In order to test this, we summarized the means for each of the main subjects included in this study: the Activity cultural dimension, the Relationship cultural dimension, choice of electives and ideal job characteristics.

ANOVA analysis showed that the means were significantly ($p < 0.000$) different across countries for all of the four subjects, with F-values ranging between 20.113 and 26.580. There was a systematic pattern in these differences with low means on all four main subjects for Sweden, Denmark, Finland, the Netherlands (both samples[6]), Austria, France and Japan and high mean scores for Greece, Turkey, Hong Kong, Malaysia, Taiwan, India, Chile and Mexico. The other countries fell between these two extremes. It is therefore apparent that a response effect is present in our data, possibly caused by an acquiescence bias for many of our Asian and Latin-American respondents, and a disacquiescence bias for our Scandinavian, many of our Northern European as well as our Japanese respondents. Since the response effect in many countries differed from the response effect in the US comparison group (see below for a discussion of this comparison group), it was necessary to remove this effect before making comparisons between the US comparison group on the one hand and individual countries on the other hand.

The established procedure for removing bias associated with scale response is within-person standardization across the instrument (Leung and Bond, 1989). However, if data are standardized with respect to the instrument as a whole, the scores for one aspect of the questionnaire affect the scores for another, reducing the validity of cross-country comparisons at the level of different aspects of the questionnaire (Maznevski et al., 2002). We therefore chose to standardize the data within-person and within-subject (Activity dimension, Relationship dimension, Electives, Ideal Job Types). A further motivation for within-subject standardisation is that for all of the four subjects we were interested in the relative importance that respondents attach to each aspect, e.g. Activity *doing* vs. Activity *being* vs. Activity *thinking*. A standardisation across the instrument as a whole would lose some of this important information. All scales were subsequently recalculated using the standardized data.

SAMPLE AND QUESTIONNAIRE ADMINISTRATION

The project coordinator recruited country collaborators through personal contacts and networking at professional conferences such as Academy of Management. Once the project had started, several researchers contacted the project coordinator directly offering to collect data in their country. All country collaborators receive a 15-page document containing very detailed instructions about the aim of the study; items and constructs; results of the pilot study; translation, data collection and data entry procedures; as well as agreements about co-authorship. All collaborators received access to the final data set. A document with personal introductions of all collaborators was prepared to promote group cohesion and facilitate networking among collaborators.

We used university students as respondents in our study. When studying culture, differences between students and other sample types, such as managers, tend to be unimportant (Triandis et al. 2001, Keating, Martin & Szabo, 2002). Hence students can be used as a good approximation of the general survey population in management studies. Although we would of course expect students to reply differently to the elective and ideal job questions than managers, we have no reason to assume that there would be a difference between students and managers in terms of their response to different lan-

guage questionnaires. Furthermore, in order to isolate the language effect from a multitude of possible intervening factors, we need samples that are strictly comparable on every aspect except for the language of the questionnaire. It is extremely difficult to realize this with a managerial sample at all, next to impossible to do this in a cross-country comparison, and completely unfeasible in a comparison across 24 countries.

Respondents were final year university students following a course in Business Administration, Business & Management, Commerce or a similar subject.[7] They were generally between 21 and 22 years old. The gender distribution varied from 27% female in India to 77% female in Hong Kong. International students were excluded from our sample, so that our comparisons only included students that could be assumed to be representative of the country they studied in. The resulting sample sizes ranged from 85 for Russia to 210 for one of the Dutch samples, but for most countries were around 100. Data were collected between March 2001 and April 2003.

Collaborators were instructed to make sure that the different language versions were randomly distributed. In most countries English and native language questionnaires were distributed in the same class. In other countries, different classes of the same subject or related subjects were used to separate English and native language questionnaires. Respondents were not allowed to choose which language version they completed and were not told about the aim of the study until after they completed the questionnaire. They were informed the study involved a comparison of values and opinions of students across countries. An equal number of English-language and native-language questionnaires were distributed.

To verify whether collaborators had succeeded in the randomization process, we tested whether there was a difference in age and gender distribution between the different language versions in each country. We also tested whether the two language groups differed systematically on the question: "How typical do you consider your view to be of people who live in the country in which you were born?" *None* of the 25 country samples showed a significant difference between the language versions on the "typical view" question, which shows that there were no systematic differences between the two lan-

guage samples. Nine of the 25 country samples showed a significant difference in age between the different language versions. However in five of these nine cases the age of the native-language sample was closer to our US comparison group than the English-language group, so that any cultural accommodation effects would be underestimated rather than inflated. In three of the four remaining cases the significant age difference was caused by a limited number of outliers (older students) in the native-language group. When these outliers were removed from the analysis, significant age differences disappeared. Only in Bulgaria did the age difference seem to be systematic and fairly large (14 months). This might have been caused by the fact that the two versions of the questionnaires were distributed in different groups. However, the two language versions did not differ in terms of their gender distribution or their response to the typical view question, so the importance of the age difference might be limited. Three of the 25 country samples showed a significant difference in gender distribution between the different language versions, but in only two of these three cases (Turkey & Finland) the gender distribution in the English-language sample was closer to our US comparison group than to the native-language group. In both countries, however, the native language sample was slightly closer to the English control sample in terms of age and in neither of the two countries did the two language samples differ in their response to the typical view question, so the importance of the gender difference might be limited. Also, in both countries both versions of the questionnaire were distributed in the same class, so any differences are probably accidental, rather than systematic. We therefore conclude that our two language samples in each country were closely matched.

To construct the Anglophone comparison group necessary to test our hypotheses, we collected data in the US, the major Anglophone country in the world. Before testing our hypotheses, we had to verify whether there were any significant "culture effects" for the variables under investigation. If there is no culture effect, than there can be no cultural accommodation effect. Using t-tests, we therefore first tested whether there were significant differences between our US comparison group and respondents in other countries that responded to a native-language questionnaire. These tests indicated that there were no significant "culture effects" for 71 of the 150 comparisons of the cultural dimensions (6

variations for 25 samples), for 36 of the 100 comparisons of the ideal job type (4 job types for 25 samples) and for 24 of the 75 comparisons of the elective questions (3 elective scales for 25 samples). So 32-46% of the variables showed no culture effect in a comparison between our US comparison group and other countries.[8] In the next section, we will use the remaining comparisons to test our hypothesis.

RESULTS

We hypothesized that when replying to a questionnaire in English non-native English speakers would culturally accommodate their answers to the values embedded in the English language. In our study, this would mean that within each country responses to the English-language questionnaire would be closer to the US comparison group than responses to the native-language questionnaire. As can be seen in Table 3, a language effect - a significant difference between the native-language sample and the English-language sample - is present for all concepts in our study. Overall, a language effect is present for 56 of the 79 (i.e. 71%) of the comparisons for cultural variations, for 25 of the 64 (i.e. 39%) of the comparisons for ideal job type and for 26 of the 51 (i.e. 51%) of the comparisons for electives. Only four of the 25 samples did not show any language effect for the cultural variations, while this was the case for seven of the 25 samples for the ideal job type and nine of the 25 samples for the electives.

Insert Table 3 about here

Of the 56 comparisons that showed a language effect for the cultural variations, 43 showed a full cultural accommodation effect, i.e. the English-language sample differed significantly from the native-language sample, but did not differ significantly from the US control sample. In addition, eight comparisons showed a partial cultural accommodation effect: responses of the English-language sample lay between the native-language sample and the US control sample, but were not significantly different from either, while the US comparison group did differ from the native-language sample. In these cases, responses to the English-language questionnaire showed some movement towards the US comparison group. Furthermore, we found four cases of cultural crossvergence: responses of the English-language

16

sample lay between the native-language sample and the US control sample, but were significantly different from both. In these cases, the English-language sample showed a unique pattern that was influenced by both the native-language values and the US values. Finally, we found one case of cultural over-accommodation: responses of the English-language sample were even more extreme than the US comparison group and all three language groups differed significantly from each other. So in total we found some evidence of a cultural accommodation effect for 71% (56 out of the 79) of the cultural variations comparisons.

Of the 25 comparisons that showed a language effect for the ideal job types, twelve showed full cultural accommodation, six showed partial cultural accommodation, three showed cultural crossvergence and four showed cultural affirmation. So in total we found some evidence of a cultural accommodation effect for around one third (21 out of the 64) of the ideal job type comparisons.

Of the 26 comparisons that showed a language effect for electives, fifteen showed full cultural accommodation, two showed partial cultural accommodation, five showed cultural over-accommodation, three showed crossvergence and one showed cultural affirmation: responses from the English-language sample lay further away from the Anglophone control sample than responses from the native-language sample, while responses for all three samples were significantly different from each other.. So in total we found some evidence for a cultural accommodation effect for nearly half (25 out of 51) of the elective comparisons.

Overall we therefore found a high level of support for our hypothesis. Although cultural accommodation was by no means universal, it occurred in a very substantial number of cases. As was expected, language and cultural accommodation effects were stronger for the cultural variations than for the ideal job type questions and the elective questions. However, even for the Ideal Job Type and elective questions, which were expected to be more neutral, a substantial number of countries showed language and cultural accommodation effects. We will come back to this phenomenon in the discussion section. First, however, we will perform a sensitivity analysis for our results.

SENSITIVITY ANALYSIS

As we have discussed in our methodology section, some of the scales used in this article did not show acceptable levels of reliability, so the language effects might simply be an artifact of inadequate scale reliability. Our results, however, do not support this assumption. For the cultural variations, the scale with the lowest reliability (Activity *being*) also showed the lowest proportion of language effects, while the proportion of language effects for the other scale with low reliability (Relationship *collectivism*) did not differ from scales with a higher reliability. For the ideal job types, the scale with the highest reliability (Ideal Job *intrinsic*) *did* show the lowest proportion of language effects. However, the three other scales also displayed a reasonable level of reliability (Cronbach's alpha: 0.61-0.64). For the elective scales, the scale with the highest level of reliability (Electives *lecturer*) actually showed the highest proportion of language effect. We therefore have no reason to suspect that the language effects we found in this study would simply be an artifact of low scale reliability.

In a study like this, the quality of the questionnaire translations is crucial. As we have argued in the methodology section, translations have been executed with great care and scale reliabilities did not differ systematically between the three language versions. In addition, the use of scales, and the fact that we were looking for overall patterns across 24 countries, is likely to mitigate the impact of any potentially remaining translation inaccuracies. Furthermore, our sample allows us to verify the accuracy of translations because it includes four sets of countries which share the same language: Austria and Germany; Spain, Chile and Mexico; Brazil and Portugal; and China, Taiwan and Hong Kong. In Austria and Germany identical questionnaires were used, while in the other countries only minimal changes were made to reflect regional language differences. If the language effects that we found in this paper were simply due to inaccurate translations, we would expect these four sets of countries to show nearly identical differences. This is generally not the case. Germany and Austria do share a language effect for "Electives *lecturer*", but in addition, Austria displays another four language effects for the cultural variations and Germany displays a language effect for "Ideal Job *balance & relationships*". Spain, Chile & Mexico share similar language effects for the Activity *doing* dimension and the Relationship *hierarchy* di-

mension, but each of these three countries also shows additional language effects that are not shared by the other countries. Brazil and Portugal share similar language effects for the Activity *thinking* dimension, but Brazil shows one and Portugal three further language effects not shared by the other country. China, Taiwan and Hong Kong do not share *any* common language effects, although China and Taiwan share one and Hong Kong and Taiwan share two. However, China and Taiwan also display additional unique language effects. Although, as discussed in the methodology section, it is difficult to guarantee translation accuracy with absolute certainty, this comparison provides us further evidence that language effects are unlikely to be due purely to inaccurate translation.

Finally, although we tried to make sure that samples across countries were broadly comparable - all respondents were students of a similar age and studying similar subjects - our samples were convenience samples and none of the universities were specifically targeted as being representative of universities in the country in question. This might mean that had we targeted other universities, the effects might have been slightly different. Since we are interested in the overall effect across 24 countries, we do not consider this to be a major problem. However, we did take the opportunity to collect data at two Dutch universities, one in Groningen, the smallish capital of the rural province Groningen in the north of the Netherlands, and one in Rotterdam, capital of South Holland, and one of the largest cities in the Netherlands. As Table 3 shows, the culture and language effects are identical for the two universities across the three electives scales. Table 3 shows the same is true for four of the six cultural variations and three of the four ideal job scales. We therefore have no reason to believe that the overall picture in terms of language effects would have been very different had we targeted different universities.

DISCUSSION

Our results showed that there were substantial language effects for all three constructs we studied: cultural dimensions (72% of the comparisons), ideal job type (39%) and reasons for choosing electives (51%). In all but five of the 107 cases that showed a language effect, these effects gave either full or partial confirmation of the cultural accommodation thesis.[9] In 69% of these cases, the convergence was

complete, i.e. the US control sample and English-language sample were not significantly different from each other, but were significantly different from the native-language sample. In a further 16% of the cases we found partial cultural accommodation: responses of the English-language sample lay between the native-language sample and the US control sample, but were not significantly different from either, while the US comparison group did differ from the native-language sample. Six percent of the cases showed cultural over-accommodation: responses of the English-language sample were even more extreme than the US comparison group and all three language groups differed significantly from each other. Finally, 10% of the cases showed crossvergence: responses of the English-language sample lay between the native-language sample and the US control sample, but were significantly different from both.

When we added up the cases where language effects were present for the three types of questions, language effects were found in all but one country (Finland) in our survey. The three types of questions did differ, however to the extent that language effects were present "across the board", i.e. in all countries. For the cultural values questions only four countries did not show *any* language effect, while for the ideal job type of questions this was true for seven of the 25 countries and for the electives for nine out of 25 countries. We therefore found some support for our assumption that cultural accommodation will be more important for the culturally-charged cultural dimension variables. However, cultural accommodation turned out to be quite important for both the culturally-intermediate Ideal-Job-Type variables and even more surprisingly for the supposedly culturally-neutral Elective variables.

Retrospectively the language effect for the Ideal-Job-Type variables might be less surprising than it would seem at first sight. Many of the ideal-job-type characteristics might be considered to have cultural elements. The Ideal-Job-Type factors *job-intrinsic* and *serving* could be interpreted as approximations of the Individualism and Collectivism dimensions, while the Ideal-Job-Type factors *balance & relationships* and *money & achievement* could be interpreted as approximations of Hofstede's Femininity/Masculinity dimensions. In fact some of the questions relating to ideal job characteristics formed the basis of Hofstede's Individualism/Collectivism and Femininity/Masculinity dimensions. For the elective

choice variables, the language effect and cultural accommodation effect was strongest for the questions related to lecturers. In contrast to most of the other elective questions,[10] these questions referred to human interaction and therefore might be considered to be more "culture-charged" and hence more susceptible to a language effect than the other questions. It is interesting to note that, overall, the proportion of language effects was highest in variables that relate to personal relationships: i.e. all cultural variations dealing with relationships, the Ideal Job Type *balance & relationships* and the Electives *lecturer* variable.

Cultural accommodation can have an important impact on conclusions drawn from cross-national studies. We will illustrate this with a country example. When using a Japanese questionnaire, a researcher comparing the Activity and Relationship cultural dimensions between Japan and the USA would conclude that there are significant differences between the two countries on three of the six variations, while the remaining three variations show some differences too. Another researcher, using the same questionnaire in English, would have to conclude that Japanese and US respondents are nearly identical on five of the six variations and show a small, non-significant difference on the sixth. When comparing Ideal Job Types, both researchers would find the two countries to be significantly ($p < 0.001$) different on all four Ideal Job Types. However, cultural accommodation would cause the magnitude of the difference for the Ideal Job Types *balance & relationships* and *money & advancement* to be underestimated with an English-language questionnaire, resulting in a lower level of significance ($p < 0.05$). Cultural over-accommodation would completely reverse a comparison of the reasons for choosing electives. A Japanese questionnaire will lead to the conclusion that Japanese respondents attach significantly less importance to intrinsic reasons and more importance to extrinsic reasons than US respondents, while an English questionnaire will find significant support for the reverse result.

Our study is not without limitations. As we have discussed before scale reliabilities were rather low for some of the scales. We cannot guarantee absolute accuracy of our translations, nor can we guarantee that different universities in the same countries would have shown similar results. However, as we have discussed in some detail in the previous section, we do not believe that any of these limita-

tions invalidates the general conclusion of our study: language and cultural accommodation effects are important and merit serious consideration in any cross-national study.

CONCLUSION

Our study has confirmed the result of earlier studies that found that language has an impact on the way people respond to questions relating to cultural values. It extends earlier studies by confirming this effect for a large number of countries and a range of concepts. Overall, our study has shown that a decision on the language of the questionnaire should be a key aspect of any cross-national study design. Where questions comprise an element of culture – and we have shown that this might be the case even with questions that at first glance would be considered neutral – the use of English-language questionnaires might obscure important differences between countries. If differences between countries are of interest in the study design, as they will be in most cross-national studies, researchers seem to have little choice but to accept the cost and inconvenience of questionnaire translation.

---

[1] Please note that this list indicates the collaborators' current affiliations, which are not always the universities at which the data were collected. Further some collaborators were not working at that particular university at the time of data collection either, but enlisted local collaborators in their country-of-origin to help with the data collection.

[2] Assessing the descriptive analysis in Ralston et al. (1995), we find that some of the language differences might have been caused by differences in other variables. Managers who responded to the English version of the questionnaire are closer to the American managers in terms of age, years employed, level of employment and size of the company than managers who responded to the Chinese version.

[3] Sample questions for the dimensions are: "People should work hard and sacrifice enjoyment to accomplish important things" (Activity *doing*), "People should take time to enjoy all aspects of life, even if it means not getting the work done" (Activity *being*), "Decisions should be made based on analysis, not intuition or emotional feelings" (Activity *thinking*), "People should satisfy their own needs before they think of others' needs" (Relationship *individualism*), "Good team members subordinate their own interests to those of the team" (Relationship *collectivism*), "People at lower levels in an organisation should not expect to have much power" (Relationship *hierarchy*).

[4] A pilot study was conducted in the UK in November 2000, where we tested different scale anchors, running from never to always, but these were not well received by the respondents.

[5] We can also use this comparison of scale reliabilities between language versions to assess whether non-native English speakers had difficulty in replying to the English-language questionnaire. If this was the case we would expect systematically lower scale reliabilities for the English-as-non-native-language group than for the English-as-native-language group. This was the case for the Activity *being* dimension and the Relationship *collectivism* dimensions, which, as we have seen above, were generally the least reliable scales. There were differences on two of the other cultural dimension scales and on the elective/lecturer scale as well, but reliabilities for the English-as-non-native-language group were still acceptable. The other eight scales did not show a lower reliability for the English-as-non-native-language group and hence we can be reasonably confident that our respondents were able to reply adequately in English.

[6] See the results section for a systematic comparison of the two Dutch samples.

[7] The Japanese sample was an exception to this. As it proved difficult to find a sample with sufficient English-language skills, we had to resort to students majoring in English. Though this might impact on our between-country comparisons, since through their exposure to the English language these students might be culturally closer to the US comparison group than Business students, it should not impact on our between-language comparisons.

8 This does not mean that there are no country differences as such on these variables, only that the difference between the US comparison group and the country in question was not significant. In fact an overall ANOVA analysis between countries showed that highly significant differences between countries were present for all variables included in this study with F-values varying from 4.748 for Ideal Job *Balance and Relationships* to 18.961 for Electives *intrinsic*.
9 The remaining five cases showed cultural affirmation.
10 The exception is the "Because my friends are choosing it".

REFERENCES

Bond, M.H. & Yang, K-S. (1982): Ethnic Affirmation versus Cross-Cultural Accommodation. The Variable Impact of Questionnaire Language on Chinese Bilinguals from Hong Kong, *Journal of Cross-Cultural Psychology*, Vol. 13, No. 2, pp. 169-185.

Botha, E. (1970): The effect of language on values expressed by bilinguals, *Journal of Social Psychology*, Vol. 80, pp. 143-145.

Brislin, R.W. (1980): Translation and content analysis of oral and written materials, in: H.C. Triandis & J.W. Berry (Eds.), *Handbook of cross-cultural psychology: Methodology*, Vol. 2, pp. 389-444, Boston, MA: Allyn and Bacon.

Brislin, R.W. (1986): The Wording and Translation of Research Instruments, in Lonner, W.J., & Berry, J.W.: *Field methods in cross-cultural research*, pp. 137-164, Beverly Hills: Sage Publications.

Candell, G.I. & Hulin, C.L. (1986): Cross-language and Cross-cultural comparisons in scale translations, *Journal of Cross-Cultural Psychology*, Vol. 17, No. 4, pp. 417-440.

Earle, M.J. (1969): A cross-cultural and cross-language comparison of dogmatism scores, *Journal of Social Psychology*, Vol. 79, pp. 19-24.

Erdener, C.B.; Wongtada, N. & Fagenson. E. (1996) The 'Dual Response' Phenomenon Among Bilingual Subjects in Management Research," *Business Research Center Working Papers Series*, WP96005, School of Business, Hong Kong Baptist University.

Harzing, A.W.K.; Maznevski, M. et al. (2002) The Interaction Between Language and Culture: A Test of the Cultural Accommodation Hypothesis in Seven Countries.

Hofstede, G. (1980): *Culture's Consequences. International Differences in Work-Related Values*, London: SAGE Publications.

Hulin, C.L. & Mayer, L.J. (1986): Psychometric Equivalence of a Translation of the Job Descriptive Index into Hebrew, *Journal of Applied Psychology*, Vol. 71, No. 1, pp. 83-94.

Keating, M.A., Martin, G.S. & Szabo, E. (2002): Do managers and students share the same perceptions of societal culture?, *International Journal of Intercultural Relations*, vol. 26, pp. 633-652.

Katerberg, R., Smith, F.J. & Hoy, S. (1977): Language, Time and Person Effects on Attitude Scale Translations, *Journal of Applied Psychology*, Vol. 62, No. 4, pp. 385-391.

Kemmelmeier, M. et al. (2003): Individualism, Collectivism and Authoritarianism in Seven Societies, *Journal of Cross-Cultural Psychology*, vol. 34, no. 3, pp. 304-322.

Kluckhohn, F.R., & Strodtbeck, F.L. (1961): *Variations in value orientations*, Evanston, Illinois: Row, Peterson & Company.

Leung, K. & Bond, M. H. (1989): On the empirical investigation of dimensions for cross-cultural comparisons, *Journal of Cross-Cultural Psychology*, Vol. 20, No. 2, pp. 133-151.

Maznevski, M.L., DiStefano, J.J., Gomez, C.B., Noorderhaven, N.G. & Wu, P.-C. (2002): Cultural Dimensions at the Individual Level of Analysis: The Cultural Orientations Framework, *The International Journal of Cross Cultural Management*, vol. 2, no. 3, pp. 275-298.

Mullen, M.R. (1995): Diagnosing Measurement Equivalence in Cross-National Research, *Journal of International Business Studies*, Vol. 26, No. 3, pp. 573-596.

Ralston, D.A., Cunniff, M.K. & Gustafson, D.J. (1995): Cultural Accommodation: The Effect of Language on the Response of Bilingual Hong Kong Chinese Managers, *Journal of Cross-Cultural Psychology*, Vol. 26, No. 6, pp. 714-727.

Sanchez, J.I.; Alonso, A.; Spector, P. (2000): Linguistic Effects in Translated Organizational Measures: A Study of Bilinguals, *Paper present at the annual Academy of Management meeting* (4-9 August), Toronto.

Schaffer, B.S. & Riordan, C.M. (2003): A Review of Cross-Cultural Methodologies for Organizational Research: A Best Practices Approach, *Organizational Research Methods*, vol. 6, no. 2, pp. 169-215.

Schemerhorn, J.R. (1990): An Empirical Reminder About Language Effects in Cross-Cultural Business and Management Research: The Case of Bilingual Subjects, *Hong Kong Journal of Business Management*, vol. VIII, pp. 57-66.

Shafiro, M.V.; Himelein, M.J. & Best, D.L. (2003): Ukrainian and U.S. American Females: Differences in Individualism/Collectivism and Gender Attitudes, *Journal of Cross-Cultural Psychology*, vol. 34, no. 3, pp. 297-303.

Singh, J. (1995): Measurement Issues in Cross-National Research, *Journal of International Business Studies*, Vol. 26, No. 3 (Third Quarter), pp. 597-620.

Sirota, D. & Greenwood, M. (1971): Understand your overseas workforce, *Harvard Business Review*, Jan-Feb., Vol. 49, No. 1, pp. 53-60.

Thomas, D.C. & Au, K. (2002) The Effect of Cultural Differences on Behavioral Responses to Low Job Satisfaction, *Journal of International Business Studies*, vol. 33, No. 2, pp. 309-326.

Triandis, H.C. et al. (2001): Culture and deception in Business Negotiations: A Multilevel Analysis, *International Journal of Cross Cultural Management*, Vol. 1, No. 1, pp. 73-90.

Tyson, G.A., Doctor, E.A. & Mentis, M. (1988): A Psycholinguistic Perspective on Bilinguals' Discrepant Questionnaire Response, *Journal of Cross-Cultural Psychology*, Vol. 19, No. 4, pp. 413-426.

Yang, K.S. & Bond, M.H. (1980): Ethnic affirmation by Chinese bilinguals, *Journal of Cross-Cultural Psychology*, Vol. 11, No. 4, pp. 411-425.

Yang, K.S. & Bond, M.H. (1990): Exploring Implicit Personality Theories With Indigenous or Imported Constructs: The Chinese Case, *Journal of Personality and Social Psychology*, Vol. 58, No. 6, pp. 1087-1095.

Usunier, J.-C. (1998): *International and Cross-Cultural Management Research*, Sage Publications: London.

Vijver, F.J.R. van de & Leung, K. (2000) Methodological Issues in Psychological Research on Culture, *Journal of Cross-Cultural Psychology*, Vol. 31, No. 1, pp. 33-51.

Table 1: Scale reliabilities of the scales used in the analysis

Scale	Reliability (Cronbach's alpha)	Scale	Reliability (Cronbach's alpha)
Culture Activity dimension		*Elective choice*	
Activity being	0.47	Intrinsic reasons	0.53
Activity doing	0.59	Extrinsic reasons	0.66
Activity thinking	0.72	Lecturer-related reasons	0.70
Culture Relationship dimension		*Ideal job types*	
Relationship individualism	0.58	Balance & relationships	0.63
Relationship collectivism	0.51	Job intrinsic	0.73
Relationship hierarchy	0.62	Money & Advancement	0.61
		Serving	0.64

Table 2: Factor analysis for ideal job questions

	Component			
	1	2	3	4
Job: Have friendly colleagues who help each other	.709			
Job: Have a good working relationship with your direct supervisor	.690			
Job: Have an opportunity to balance your work and private life	.643			
Job: Have security of employment	.582		.424	
Job: Have little tension and stress on the job	.490			
Job: Have the opportunity to take full responsibility for a task		-.682		
Job: Make a real contribution to the success of your organisation		-.674		
Job: Have challenging work to do		-.643		
Job: Have an element of variety and adventure in the job		-.634		
Job: Have considerable freedom to adapt your own approach to the job		-.569		
Job: Be consulted by your direct superior in his/her decisions		-.540		
Job: Have an opportunity for high earnings			.778	
Job: Work in a prestigious, successful comparny or organisation			.694	
Job: Have an opportunity for advancement to higher level jobs		-.431	.674	
Job: Serve your country				-.764
Job: Have an opportunity for helping other people				-.700
Job: Work according to clear and stable rules and regulations				-.691
Job: Have the opportunity to share responsibility for a task with others	.458			-.494

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

Table 3: Language effects in culture, ideal job and elective questions

Country	Activity Being	Activity Doing	Activity Thinking	Relationsh. Collectivism	Relationsh. Hierarchy	Relationsh. Individualism	IJ: Balance & Relationsh.	IJ: Intrinsic	IJ: Money & Advance	IJ: Serving	Electives Intrinsic	Electives Lecturer	Electives Extrinsic
Austria		AC†	AC	AC†	--	AC	--	--		--	--	AC†	--
Brazil			(AC)		--	AC†	--	--	--	--	--	--	--
Bulgaria	AC†	--		<AC>	--	AC†		--		<AC>	--	[AC]	AF
Chile	AC†	AC†			AC	AC	(AC)†	--	--	--			
China	--		--				AC†		--		(AC)	AC	
Denmark		AC†	(AC)	AC†		AC		--		--	--	AC	--
Finland	--		--				--	--		--	--		--
France	AC	(AC)		AC†		AC	--	--		AF	--	--	--
Germany							(AC)	--		--	--	(AC)	--
Greece	AC				--		--	--		--			
Hong Kong				AC		AC	--					--	--
India	<AC>†	(AC)	(AC)					--	AC†	--		--	--
Japan			AC	AC	AC		--	--	(AC)	AF	[AC]		[AC]
Lithuania	--	--		--		--		--		--	AC†	[AC]	AC
Malaysia	--	--	<AC>					(AC)		AC		--	
Mexico	--	AC†			AC	AC		<AC>	(AC)	AC	--		AC
Netherlands (1)	--	(AC)	--			AC	AC†	--			<AC>		AC
Netherlands (2)	AC	AC				AC†	AC	--		AF	<AC>†		AC
Poland			[AC]	AC		AC		<AC>	AC	AC		AC	--
Portugal			(AC)	AC†	(AC)			--	--	--	--	AC†	
Russia				--		--	AC†	--		--	AC†		--
Spain	--	(AC)		(AC)	AC		AC†	AC				--	
Sweden	--	AC†	--				AC†					--	--
Taiwan				AC†	AC	AC		--		AF	[AC]		AC
Turkey	AC	AC		--	AC	<AC>		(AC)			AC	AC†	AC
% language effects	7/15: 47%	11/13: 85%	8/11: 73%	10/13: 77%	7/11: 64%	13/15: 87%	8/14: 57%	5/23: 22%	4/8: 50%	8/19: 42%	8/18: 44%	13/17: 76%	5/16: 31%

Summarised % of language effects: Cultural dimensions : 72%, Ideal Job dimensions : 39%, Electives: 51%
Countries without any language effect: Cultural dimensions : 4/25, Ideal Job dimensions : 7/25, Electives 9/25

Legend for table 3

N = Native-language sample, E = English-language sample, U = US comparison group
= no significant difference between groups, ≠ significant difference between groups

	=	no culture effect (N = U, p < 0.05, 2-tailed)
--	=	no language effect (N+E ≠ U, p < 0.05, 2-tailed)
AC	=	full cultural accommodation (N ≠ E+U, p < 0.05, 2-tailed)
AC†	=	full cultural accommodation (N ≠ E+U , p < 0.10, 2-tailed)
(AC)	=	partial cultural accommodation (N = E & E = U, but N ≠ U and E lies between N & A, p < 0.05, 2-tailed)
(AC)†	=	partial cultural accommodation (N = E & E = U, but N ≠ U and E lies between N & A, p < 0.10, 2-tailed)
[AC]	=	cultural over-accommodation (N ≠ U ≠ E, p < 0.05, 2-tailed)
<AC>	=	cultural crossvergence (N ≠ E ≠ U, p < 0.05, 2-tailed)
<AC>†	=	cultural crossvergence (N ≠ E ≠ U, p < 0.10, 2-tailed)
AF	=	cultural affirmation (E ≠ N ≠ U, p < 0.05, 2-tailed)

27