Challenges in international survey research: a review with illustrations and suggested solutions for best practice

Anne-Wil Harzing*

Department of Management and Marketing,
University of Melbourne,
Parkville Campus, Victoria 3010, Australia
Email: harzing@unimelb.edu.au
*Corresponding author

B. Sebastian Reiche

IESE Business School,
Department of Managing People in Organizations,
Ave. Pearson, 21, Barcelona 08034, Spain
Email: sreiche@iese.edu

Markus Pudelko

Department of International Business,
Tübingen University,
Melanchthonstraße 30, 72074 Tübingen, Germany
Email: markus.pudelko@uni-tuebingen.de

Abstract: When conducting international research projects, scholars face a myriad of challenges that reach beyond those encountered in domestic research. In this paper, we explore the specific issues related to international survey research, focusing on the different stages of the research process that include defining the study population and gaining data access, survey development, data collection, data analysis, and finally publication of the results. For each stage, we review the pertinent literature, provide illustrations based on examples from our own research projects, and offer possible solutions to address the inherent challenges by formulating suggestions for improving the quality of international survey research.

Keywords: international survey research; cross-cultural research; equivalence; best practice.

Reference to this paper should be made as follows: Harzing, A-W., Reiche, B.S. and Pudelko, M. (2013) 'Challenges in international survey research: a review with illustrations and suggested solutions for best practice', *European J. International Management*, Vol. 7, No. 1, pp.112–134.

Biographical notes: Anne-Wil Harzing is Professor in International Management at the University of Melbourne, Australia. Her research interests include international HRM, expatriate management, HQ-subsidiary relationships, the role of language in international business, and the quality and

Copyright © 2013 Inderscience Enterprises Ltd.

impact of research. Since 1999 she has maintained an extensive website (www.harzing.com) with resources for international and cross-cultural management as well as academic publishing and bibliometrics. Her latest book is *The Publish or Perish Book: your guide to effective and responsible citation analysis*.

B. Sebastian Reiche is Assistant Professor in the Department of Managing People in Organizations at IESE Business School, Spain. He earned his PhD from the University of Melbourne, Australia. His research focuses on international assignments and global work, cross-cultural management, knowledge transfer, social capital and employee retention, and has appeared in a number of scholarly outlets including the *Journal of International Business Studies*, *Human Resource Management* and *Journal of World Business*. He regularly blogs on the topic of expatriation (www.blog.iese.edu/expatriatus).

Markus Pudelko is Professor of International Business at Tübingen University School of Business and Economics and worked previously for the University of Edinburgh Business School. He has earned Masters degrees in Business Studies (University of Cologne), Economics (Sorbonne University) and International Management (Community of European Management Schools – CEMS) and a PhD (University of Cologne). His current research is on HQ-subsidiary relationships, comparative HRM, Japanese management, multinational teams, the impact of language on international business, and cross-cultural management.

1 Introduction

In contrast to what the wealth of textbooks on conducting empirical research seems to suggest, the actual research process is quite messy in nature. In fact, it can be viewed "as a set of dilemmas to be 'lived with'; and […] as an effort to keep from becoming impaled on one or another horn of one or more of these dilemmas" (McGrath, 1982, p.69). From this perspective, embarking on a cross-national research project introduces many additional dilemmas.

In this paper, we explore the challenges related to doing international survey research and illustrate them with examples from our own research projects. In addition, we provide possible solutions to address them by formulating suggestions for improving the quality of international survey research. Whilst most of our suggestions are targeted at researchers, some of our observations are targeted at other constituents of international business research, such as reviewers, journal editors or business schools employing international business researchers.

Whereas cross-cultural investigation is not limited to survey research and includes a range of qualitative methods of data collection (see Marschan-Piekkari and Welch, 2004 for a good overview), we focus our discussion on the collection of international and cross-cultural data through questionnaires. In the following sections, we will discuss the various methodological challenges in more detail. We will structure our discussion along the various stages of the research project, referring to study population and data access, survey development, data collection, data analysis, and finally publication of the results. For each of these topics, we will provide a review of the published literature in the field as well as illustrations based on our own experience.

Our illustrations are based on three large scale international survey projects we have been involved with in the last five years. In the first project, we studied work values and leadership styles across twenty-two countries. However, the primary aim of the study was to assess the impact of the language of the questionnaire on the responses provided. Data were collected by local country collaborators using classes of MBA students as respondents. The second project involved a study on the role of language in HQ-subsidiary relationships. Data were collected through paper and online surveys in twelve countries. Although data were largely collected locally, the project was centrally controlled and most of the design and data collection was executed by only two researchers. The third project dealt with the study of trust in manager-subordinate relationships across cultures and involved data collection in 18 countries. The overall project was led by three researchers; yet actual data collection was decentralised to country collaborators. Data were collected through both paper and online surveys from managers and up to three of their subordinates.

2 Study population and data access

The survey population is a crucial concept in empirical research as it determines the set of entities from which the sample can be drawn and affects both the internal and external validity of a study's results. Internal validity refers to the extent to which the manipulation of an independent variable is the sole cause of change in a dependent variable. In contrast, external validity concerns the generalisability of the results.

Internal validity is threatened if the observed results are influenced by the confounding effects of extraneous variables. To control for possible extraneous variation, it is important to select a homogenous population. In this vein, Sekaran (1983) highlights the use of matched samples that are functionally equivalent across the countries of interest. For example, in international assignment research this may entail the focus on one particular group of assignees (e.g., inpatriates or short-term assignees) from different countries-of-origin. At the organisational level, this may involve limiting the analysis to MNCs of comparable size, industry affiliation or internationalisation experience. External validity is at risk if the selected sample fails to adequately represent the larger population. Using a stratified random sample can mitigate this risk, for example by ensuring relative representation of respondents across different social classes or religious groups in each country under study (e.g., Tsui et al., 2007).

Any research project is also dependent on access to sufficient data to address the research question(s) of interest. In an international research context, data access concerns not only securing an appropriate sample, but also ensuring that all data can be feasibly collected given the additional cost that are involved in cross-border mail, telephone and fax correspondence. A systematic way to identify all organisations that form part of the target population is the use of databases with information on company profiles and respective contact details (e.g., Hoover's Handbook of World Business) in combination with local address books. Commercial organisations such as D&B are also able to supply a list of addresses conforming to a specific set of requirements against a fee (see illustration 2). Furthermore, international professional organisations or consulting firms can be contacted, which may be willing to share their member or client databases in return for access to research results. However, compiling a comprehensive database of

for instance multinational corporations is not an easy task and researchers using only a single database might well find that more than half of their questionnaires do not reach the target population. This puts the usually low response rates (see illustration 7) in a slightly less negative light.

In most cases it is appropriate to address a request letter or email to either the managing director or a functional director or, if individual employees serve as the primary unit of analysis, the HR director. As managing directors rotate frequently, it is advisable to confirm the personal details prior to sending out the request (Harzing, 1999). Alternatively, one can simply address the survey to "The Managing Director" or "The HRM Director". However, this is likely to further reduce the already low response rates in "cold call" international mail surveys.

The inclusion of local collaborators in the countries of interest not only serves as an additional means to gain access to local companies but, importantly, also helps to manage the international data collection process (Harzing et al., 2005). Indeed, local collaborators can collect the data on-site and return the responses in one batch, thereby facilitating the data transmission. Also, they provide additional credibility to the research project in the local context which may help to increase response rates. Local collaborators may also provide relevant input for creating a representative sample (Tsui et al., 2007) and help with the interpretation of culture-specific findings (Harpaz, 2003).

Illustration 1: Equivalent samples across countries to ensure internal validity

In project 1, data were collected from MBA students in twenty-two different countries. The use of student samples poses limitations in terms of representativeness, and journals sometimes pose restrictions on student samples or provide guidelines under which circumstances they are accepted or not (e.g., Bello et al., 2009). Especially in developing countries students might be different from the population as a whole and might be more westernised than non-students. The same is the case for employees of foreign (Western) subsidiaries (Caprar, 2011). However, this does mean that any cross-country differences might be attenuated, so that the study provides a more stringent test of these differences. In this project we focused on MBA students as opposed to the UG student population that is frequently used in cross-cultural studies. We purposefully sampled respondents with managerial work experience, so that in their responses to our survey they could draw on real life experiences. This mitigates the traditional disadvantages of experimental designs.

One of the aims of this project was to assess the influence of the language of the questionnaire. In each country, half of the respondents completed questionnaires in English and half in their native language; the language versions were distributed randomly. To verify whether collaborators had succeeded in the randomisation process and thus ensure internal validity, we tested whether the two language groups differed on one of the background questions: "How similar are your norms and values to the majority of people in your birth country?" None of the countries in the study showed a significant difference between the language versions on this question. However, in some of the countries there was a difference in age and gender distribution between the different language versions. We therefore included age and gender as control variables in our statistical analysis. As a result we were able to use this project to write up papers relating to the actual content of the surveys, as well about language and response bias.

Illustration 2: Buying addresses from a commercial provider to gain access to data

In project 2, we aimed to study MNC subsidiaries in twelve countries, representing a variety of language and cultural contexts. It would have been very difficult and time-consuming to collate address lists from publicly available sources in each of these twelve countries, especially as the project was largely carried out by only two researchers. After a thorough investigation of the available options we therefore decided to buy a customised address list from D&B. This turned out to be easier said than done.

It took one meeting, half a dozen long phone calls and an exchange of well over 100 emails to get (or rather buy) the appropriate address lists. One reason for this was that all communication had to go through the Australian customer relations contact, who subsequently contacted the D&B International Office (DBI, located in the UK). DBI then contacted the relevant country offices for further information. Needless to say, this did not result in very prompt response to our questions. There were also several hick-ups along the way.

First, addresses in Korea, Japan and China were delivered in roman script rather than in character-based scripts. This was not a problem in Japan or Korea, although it might well have led to lower response rates. However, for China we found – after preparing the entire mailing – that letters within China can only be mailed with addresses in Chinese characters. Hence we had to go through the time-consuming and costly process of acquiring addresses in Chinese characters from the Shanghai D&B office. Second, the type of data collected by the D&B local offices varies by country. In Japan for instance, employee size is not a required field. Hence, when we requested foreign-owned subsidiaries with more than 50 employees, we received a very low number of addresses for Japan as all subsidiaries with missing data had simply been excluded. We only discovered this after a long exchange with the Japanese office querying the unexpectedly small number of records. Unfortunately, the additional Japanese records then still had to be purchased locally in Australia, resulting in further delays.

All in all it took more than three months before we had access to the final list of addresses for all countries in our study. As data in China and Japan were collected by a member of the research team who had planned only one-month stints in each of the countries, it was a stressful race against the clock. Hence researchers would be well-advised to not leave acquisition of addresses until the last minute. In total we paid approximately $A 11,000 for just over 12,000 addresses (with a free update to compensate for all the problems experienced). Of these more than 20% turned out to be unusable, because they were undeliverable, duplicate addresses, the companies were not foreign-owned, or they had folded when we used our free refresh six months later. Therefore, for researchers on a tight budget, it might pay to use public sources instead.

Whereas these problems and our solutions might be quite specific to the project concerned, they illustrate the sort of unexpected problems that can occur, in spite of very careful planning. Allowing for sufficient monetary resources and time to obtain access to data is therefore of paramount importance.

Summary of suggestions

For the first stage of international survey research – determining the study population and gaining data access – we recommend using matched samples in combination with appropriate control variables to ensure internal validity, developing a stratified random sample to ensure external validity, accessing multiple databases with information on

company profiles and respective contact details to construct the target sample, allowing sufficient time to obtain and verify these contact details, and involving local collaborators.

3 Survey development

When developing a survey, three methodological issues require special attention in an international research context: (1) choice of survey type, (2) item generation, and (3) survey language(s).

3.1 Survey type

Surveys can be conducted by face-to-face interview, telephone, fax, mail and internet. In the case of large-scale international survey research, both face-to-face and telephone interviews are usually not feasible in terms of language difficulties and the costs involved. Additionally, substantial time-zone differences between the researcher's and the respondents' location limit the available time frame for scheduling and conducting telephone conversations. The scarce research assessing the effectiveness of fax surveys (Tse and Ching, 1994) indicates that their response rates tend to be lower than traditional mail surveys and declining fax usage means they are even less likely to be effective these days. These limitations have led the majority of international researchers to rely on either paper-and-pencil surveys administered by postal mail or internet surveys.

Traditional mail surveys are not without problems, however. Mailing times can be substantial, thus delaying the data collection process. Also, the use of international prepaid postal coupons adds significant extra costs to the research project and there is some concern about the reliability of postal services in less developed countries (Harpaz, 2003). International reply-paid numbers are a more cost-effective alternative, but in many countries there are only a few post offices that offer this service. Therefore, researchers increasingly emphasise the use of email and internet-based surveys as an effective alternative (Dillman, 2006; Hewson et al., 2003). Generally, surveys administered via the internet offer several advantages over paper-and-pencil surveys. For example, internet-based questionnaire distribution involves lower cost as well as higher transmission and response speed, which is of particular importance in an international research context. This is especially true if invitations to participate in the survey are sent by email. In addition, web-based surveying entails time and cost savings with regard to data entry and reduces the risk of data entry errors as respondent data can be automatically transformed into a format ready for analysis (Hewson et al., 2003).

However, potential technical problems in internet surveys should not be underestimated. Researchers need to ensure questionnaires can be read on a variety of screen sizes and in a variety of different internet browsers. Designing a survey in different languages can be challenging as different languages need different amounts of physical space. Frequent checking and double-checking by native speakers is required. Imagine for instance having to decide whether or not Chinese characters in a scale anchor can be spread over two lines if you have no idea what they mean, or even whether they represent one word or several!

An abundant literature exists on comparing response rates of internet surveys to traditional mail surveys and several meta-analyses of online survey response rates have been conducted (e.g., Fricker and Schonlau, 2002). Although results are far from consistent, the overall conclusion is that online response rates are lower than response rates to mail surveys and that response rates have declined over time (see also the section on survey process and response rates). Two factors were most important in explaining response rates: the population studied and the number of contacts (i.e. whether or not an announcement and/or one or more reminders are sent). Finally, when offered the choice most respondents preferred mail surveys (Fricker and Schonlau, 2002).

Unfortunately, most studies on this topic are fairly old. This is problematic, given the rapid development of email and internet technology. Increasing concerns about spam in recent years might have made respondents more hesitant to respond to email invitations. Also, whereas a mail survey usually attracts at least a glance and has to be physically destroyed, email surveys and email invitations for internet surveys can be deleted without being read, and with no more than a mouse click. Finally, many countries now have anti-spam regulation that prevents the use of mass e-mailings. In the D&B illustration above, the company was unable to provide email addresses of individual officers for this reason. The increasing incidence of internet fraud as well as the misuse of information provided online might also have made respondents more hesitant to complete online surveys. Given the challenges of achieving high response in international surveys, especially with managerial populations, a mixed approach (see illustration 3 below) might offer the best result.

Illustration 3: Survey type: online versus paper surveys

In project 2, we initially intended to collect data with internet-based surveys only, with an invitation letter sent by mail. A sophisticated online survey was designed and offered in the local language in most countries. It took nearly three months to design the original (English-language) survey, after which translations were imported and their on-screen appearance fine-tuned. The whole process was much more time-consuming and costly than creating a paper survey and frequent consultation with technical experts was required. However, we expected the time and cost disadvantage to be compensated by lower distribution costs and a faster response.

Unfortunately, response rates in the first mailing in Australia/New Zealand and the Asian countries (China, Japan) were very disappointing. We therefore provided a paper option in the reminder for Australia/New Zealand and Japan. In China, no paper versions were mailed because of logistical reasons; we had already prepared the reminder before we decided to use paper surveys. In Korea, data collection was outsourced to a survey company which called potential respondents in order to boost the response rate for the small population of subsidiaries in this country. Hence, most responses in Korea took place online.

In Australia/New Zealand 70% of the responses in the reminder (where respondents were offered the choice between completing the online or paper version) were paper versions, even though this involved the additional effort of mailing the questionnaire back to the researchers. The response rate in the reminder was also more than double the response rate in the initial mailing, leading us to strongly suspect that offering a paper questionnaire positively influenced response rates. In Japan, only 25% of the responses in the reminder were paper versions. This might be partially caused by the fact that we had photocopying problems and the paper version looked less attractive than in other

countries. Even so, the response rate in the reminder was 50% higher than in the original mailing. Hence, the respondents returning the paper version might not have responded online, and even those that responded online might have appreciated the choice between a paper and online version.

In all European countries (UK, Spain, France, Germany, Sweden, Norway, Denmark, Finland), paper questionnaires were included with both the initial mailing and the reminder. In nearly all of those countries, Spain being the only exception, we received substantially more paper than online responses. Overall, nearly two third of the responses in the European countries were paper versions, running from 40% in Spain to 80% in the UK. In these countries, the response rates in the reminder were less than half than the response rates in the initial mailing, which is a fairly normal pattern in surveys (Harzing, 1999). Hence, we can be fairly certain that at least for our audience, HR managers in multinational companies, paper questionnaires were generally preferred over internet questionnaires.

A specific technical problem was encountered in China. As indicated before, we had localised questionnaires through translation; even the title of the web page was translated. When we received virtually no responses (out of more than 3000 invitation letters mailed), we asked several friends in China to double-check whether the website displayed without problems. It turned out that although the website could be accessed outside China, it could, after having been online for a while without problems, not be accessed within China anymore and just showed a blank page. After further investigation it turned out that websites hosted outside China with a Chinese webpage title were blocked by internet censoring. When we changed the webpage title to English (whilst keeping the rest of the survey the same), the website displayed without problems in China. So it is essential have someone trusted check website access in the country in question.

3.2 Item generation

Questionnaire design involves the decision about which items will best reflect the underlying construct the researcher wishes to measure. Although a wealth of existing scales is available for measuring constructs in the management discipline, these scales may not be easily transferable to a different cultural context. Implicit to this argument is the issue of construct equivalence in cross-cultural research referred to earlier (see Hult et al., 2008 for a recent review on construct equivalence in cross-cultural international business research).

In general, whether construct equivalence can be established is contingent upon the type of perspective the researcher takes towards the study of culture, namely emic or etic. The emic approach emphasises the intrinsic cultural distinctions that are meaningful to the members of a given society, whereas the etic perspective attempts to derive commonalities between cultures. Therefore, when the research project follows an emic approach, it will be restricted to uni-cultural or polycentric inquiry (Peng et al., 1991). Ethnographic studies serve as a key method to address such research issues. In contrast, survey research is primarily useful for etic considerations as it allows for cross-cultural comparisons.

Even in the case of an etic research perspective, establishing construct equivalence encompasses various difficulties. For example, Adler et al. (1989) failed to validly and reliably describe management behaviour in China as some of their measurement items

contained the Western notion of 'truth' which has different connotations in Confucian philosophy. Thus, a construct can only be meaningfully measured across cultures if it is based on a universally applicable concept in these cultures, that is, is conceptually equivalent. In this regard, whereas questionnaire translation (see next section) is necessary to 'clarify' construct elements in the local language and frame of reference it is an insufficient condition for establishing construct equivalence (Peng et al., 1991). Instead, in many cases the original scale will need to be re-constructed and existing items complemented with additional questions to appropriately capture the underlying construct. Again, the use of multinational research teams whose members are familiar with the respective local cultures may help to overcome problems related to adapting measurement scales (Harpaz, 2003). Key to a meaningful modification of existing measurement scales is a sound process of scale development (see Hinkin, 1995 for a good overview on scale development practices). However, country-specific adaptation by necessity reduces cross-cultural comparability. The feasibility of modifying existing scales to accommodate for cultural specificities is therefore limited if data from a larger amount of countries are to be compared. In that case, the collection of qualitative data, for example through interviews, can compensate for the inherent limitations of survey data. Triangulation can thus increase the robustness of the data gathered.

Illustration 4: Construct equivalence in measuring trust across cultures

Project 3 investigated the antecedents and outcomes of trust in manager-subordinate relationships, a concept whose meaning could be expected to differ across the wide variety of cultures such as Pakistan, Ivory Coast or Peru that were included in our study. To achieve construct equivalence, we employed a three-pronged approach. First, as we were interested in examining the universality of the hypothesised relationships between manager-subordinate trust and other constructs (e.g., organisational citizenship behaviour) across different cultures, we used the same measures in all countries rather than adapting our scales to each local context. Whereas scale adaptation would have been preferable for studying trust in one or a few specific cultural contexts, using 18 different trust scales would have made the cross-cultural comparability of our hypothesised relationships difficult, if not impossible. Therefore, we conducted a series of measurement invariance tests with our data set (see Data Analysis and Publication). These analyses led us to drop a few items of our original scales that were not found to be equivalent across cultures. For our measure of managerial trustworthy behaviour, we had to delete an entire sub-dimension (delegation) that failed to show cross-cultural equivalence, suggesting that this dimension was culturally more distinct.

In a second step, each local collaborator conducted between five to ten personal interviews with representatives of the local culture. In these interviews, participants were presented with a list of values that were created by the whole research team and intended to reflect the concept of trust. Participants were then asked to rate the extent to which each value would be associated with trust in their culture, explain their choice, and add additional values that they thought would be missing on the list. The interview ended with a broader discussion of the concept of trust. These qualitative data provided us with a much deeper understanding of the emic characteristics of the trust concept in each culture and served as a basis for the development of culture-specific items of trust.

Third, we conducted focus group interviews in each country to contextualise and make sense of the findings from our quantitative study. Specifically, we conducted the

following three 1-hour focus group sessions, led by the respective local collaborator: (1) a panel of three to six managers, (2) a separate panel of three to six subordinates (one per manager that participated in the manager panel), and (3) a joint panel with both managers and their subordinates. Although the managers and subordinates were different from those that participated in our survey, we were careful to match their characteristics to those of our quantitative study. These data gave us further insight into why our hypothesised relationships varied across different cultural contexts.

3.3 Survey language

The choice of survey language should be primarily determined by respondents' language proficiencies. In the case of surveying MNCs' managerial employees who are likely to possess a sufficient level of English and have been exposed to similar tertiary education in business schools around the world, the use of single-language surveys in English may be adequate. However, research has also shown that the use of English-language questionnaires might create a language bias. Important differences between countries are obscured through reduced variance in responses between countries, caused by cultural accommodation (Harzing et al., 2005) or by a lack of the respondent's confidence in responding in a non-native language (Harzing, 2006). Further research even showed that the language of the questionnaire can impact not just attitudes, but also behaviours (Akkermans et al., 2010). Especially if both native and non-native English speakers are included in an international survey, survey translation into the respective local language appears crucial. Finally, translation might also have a symbolic effect. Even if respondents are comfortable with English, translation demonstrates to the respondent that the researcher has gone through the effort and expense to make responding as easy as possible. This might well influence response rates.

As many concepts and terms entail culture-specific connotations, their mere direct translation is unlikely to transport the intended meaning. For example, the concept of feedback differs substantially across cultures. Whereas it is usually viewed as a direct, open and formalised process in the USA or the UK, many Asian countries regard feedback as a more indirect, anonymous and informal procedure (Hofstede, 1998). Without clearly specifying the intended meaning of the concept in the translated questionnaire, the researcher risks introducing systematic bias. A meaningful translation of the original version of the questionnaire requires a researcher not only to ensure overall conceptual equivalence but also to consider vocabulary, idiomatic and syntactical equivalence (Sekaran, 1983). In this vein, Brislin (1980) has suggested to use simple sentence structures as well as clear and familiar wording as much as possible to facilitate translation. In addition, by adding redundancy and necessary context for difficult phrases, the researcher is able to clarify the intended meaning.

The most frequently employed translation technique is back-translation (Brislin, 1970). In this procedure, the original version of the questionnaire is translated into the target language and subsequently translated back into the source language by a second bilingual person. The use of two independent translators increases the chances that the original meaning has been retained, ensures literal accuracy and helps to detect mistakes. However, given the earlier notion that corresponding concepts may not always exist in another language, back-translation does not guarantee overall conceptual equivalence (Peng et al., 1991). Furthermore, the more the translation is adapted to the specific local context (emic perspective), the more the comparability between countries and questionnaire versions is compromised (etic perspective).

Harpaz (2003) identifies two additional translation techniques: bilingual method and committee procedure. The former approach involves sending the original and the translated questionnaire to bilingual individuals and subsequently correcting items based on inconsistencies in their responses. In contrast, in the latter approach a committee consisting of bilingual individuals translates the questionnaire jointly and discusses possible mistakes or difficulties. Finally, to cross-check for possible translation mistakes and to ensure comprehension of the translated questionnaire among respondents, pilot-testing is particularly important in international research.

Illustration 5: Translation through committee procedures

In project 2, the study on the role of language in HQ-subsidiary relationships, data were collected in twelve countries. Two of these countries were native English speaking (UK, Australia). In the four Nordic countries (Sweden, Norway, Denmark, Finland), we expected the English language capacity of our respondents to be high enough to provide reliable responses in English. The relatively small sample sizes in these countries also meant that translation into an additional four languages was not cost effective. However, we did translate the survey instrument into Chinese, Japanese, Korean, German, French and Spanish as we could not expect all of our respondents – local HR managers – to be fluent in English.

We followed a variant of the committee procedure. The initial translation took place by bilingual research assistants under the supervision of the project coordinator. Subsequently, a focus group consisting of both the translator and two or three other bilingual students discussed the translated questionnaire in the presence of the project coordinator. First, the bilingual students were asked to carefully review the translated questionnaire one item at a time and indicate whether the text sounded "natural" to them. Subsequently, they were asked to review the original English sentence and assess whether it was equivalent to the corresponding native version. Even if only one of the students felt the items were not fully equivalent, the translator initiated discussion between the participants to find a better translation. The project coordinator was available to provide feedback on the meaning behind the questions where necessary. This review process took at least 3 hours, but for the Asian languages it usually took three sessions for each language, lasting up to 8 hours in total.

In some instances, we had to extend the committee procedure even for the accompanying cover letter. Initially, we asked a fully bilingual Japanese undergraduate student to translate this cover letter for us. However, formal "business Japanese" is so different from "regular" Japanese that Japanese only learn this specific writing style once they enter the corporate world. Consequently, we made sure that in all Asian countries, we had the letter written or at least proof-read and corrected by assistants with management experience.

Summary of suggestions

We recommend the following best practices in the development of international surveys. In terms of the choice of survey type, it is advisable to obtain help from native speakers in designing the survey, ask locals to check and test the survey, and combine both paper-and-pencil and online surveys to increase response rates. To generate survey items, it is important to first decide whether the research project is emic or etic in nature. Whereas the former case requires the adaptation of existing and development of new items, in the

latter case this may be unfeasible when interested in cross-cultural comparability. Instead, researchers should collect additional qualitative data to triangulate the study results. Finally, several translation techniques exist to adapt the survey to the local language of the countries it is diffused to. Simple sentence structures should be used to facilitate translation, and additional clarifications provided to better convey the intended meaning of certain questions. It is also important to pilot-test the survey in each country.

4 Survey process and response rates

Similar to survey development, the survey and data collection process is also likely to require substantially more time than in domestic research, as it has to be adapted to local circumstances. This is not least due to the need to manage different language versions of the questionnaire, coordinate with country collaborators and, in some cases, even employ different means of survey administration to accommodate respondents' different levels of technological proficiency. In addition, ideal times for distributing the survey may vary across countries. For example, countries have different public holidays, different peak holiday periods and even differ in terms of their end of financial year dates, which usually correspond to an increased workload for employees. However, the timing of data collection not only affects its overall length but can also influence the results. Research, for instance, has shown that the September 11 attacks had an impact on cultural values and the level of cosmopolitanism of US university students (Olivas-Luján et al., 2004).

A key challenge in any survey research is to maximise the study's response rate. Overall response rates have been found to differ significantly, both across different professions and occupational groups as well as across countries. For example, evidence suggests that response rates of managerial employees are lower than those of non-managerial staff (Baruch, 1999). In a recent meta-analysis, Cycyota and Harrison (2006) identified an overall top manager response rate of 32%. In an international research context, these rates are likely to represent an unrealistic dream. Drawing on studies conducted between 1988 and 1994, Harzing (1997) reported typical response rates for "cold call" international mail surveys to lie between 6% and 16%. Sheehan (2001) also reported declining survey response rates in the USA (both for mail and e-mail surveys). Our collective experience in supervising PhD projects in different countries suggest that the situation has not improved since then and that double-digit response rates are increasingly difficult to achieve. In addition, research has identified considerable cross-national differences that are partly contingent upon the researcher's origin. Harzing (2000), for instance, showed that higher response rates were achieved when respondents were geographically and culturally closer to the research project's originating country, were more internationally oriented and came from countries with a lower level of power distance. This home-country effect could be moderated to some extent by sending questionnaires locally rather than from one central location (see illustration 6).

Several factors have been found to influence response rates in domestic research, which exert differential effects across cultural research contexts. We will discuss three categories of strategies to increase response rates: strategies related to the questionnaire design, the survey process and incentives offered (Dillman, 2006). First, as survey appearance is a widely accepted determinant of response rates, questionnaires should be user-friendly and have a professional layout. It is also important to personalise the correspondence with potential respondents, by using real signatures and addressing

respondents individually. However, whether a survey appears well-designed to the individual respondent is highly subjective and may vary considerably across cultures: certain colours and pictures used on the cover page or throughout the survey can have culture-specific connotations, which may require slight adaptations of the survey design. Again, country collaborators and pre-tests with individuals from the target culture may facilitate this process. In addition, overall questionnaire length is considered an important predictor of response rates (e.g., Tomaskovich-Devey et al., 1994), yet may vary considerably across different languages. A questionnaire translated from its original English version into German or Finnish can be substantially longer, thereby affecting respondents' decision whether or not to complete the survey. By contrast, questionnaires translated into a language using Chinese characters can be significantly shorter. Before making a final decision about the number of measurement scales to include and thus the overall survey length, the original version should be translated into all required languages first. It is also important to note that due to respondents' different levels of language proficiency and general educational background, the average time to complete a survey may vary. It is therefore recommendable to provide respondents with a range rather than a specific estimate for the survey completion time.

Second, there are various strategies to increase response rates that concern the actual survey process. In general, it is beneficial to follow a multi-stage survey process that includes the circulation of an announcement letter and the distribution of reminders (Dillman, 2006). In addition to the actual questionnaire, these may also need to be translated into the local language. In the case of using single-language surveys, it is at the very least necessary to include a note in the local language in case the survey is forwarded by a colleague or secretary. It is particularly important to seek sponsorship for the study given the geographical and cultural distance between the researcher and the respondents. Sponsorship can be provided by an international or local professional organisation, a leading local business school, through an international committee of recommendations that includes local representatives from every target country (Harzing, 1999) or, at the level of the individual unit of analysis, of the respective participating organisations. It often takes the form of an explicit letter of endorsement that can be attached to the actual cover letter, expressing support for the study and asking for participation.

Third, incentives may be used to increase survey response rates. In an international research context, the inclusion of financial tokens, which have been shown to increase response rates (Dillman, 2006), is difficult to administer, due to currency differences, purchase power differences as well as possible differences in ethical perceptions. From this perspective, non-financial incentives may be preferable. This may entail the inclusion of a 'Thank you' note in the reminder letters, thereby thanking those who have already completed the survey. Also, promising respondents to provide them with a summary report of the overall research results and recommendations of the study is beneficial. Organisations may be particularly interested in benchmarking themselves against other firms. Again, local adaptation is likely to result in an increased effect. The main conclusion is therefore that to achieve the best results researchers should be willing to incur the additional time (and sometimes cost) to tailor data collection procedures by country.

Illustration 6: Adaptation to local circumstances by variance in data collection across countries

Project 1 collected data in class with MBA students, hence applying fairly standardised data collection procedures. Even so, local collaborators had complete freedom to introduce the project in a way they saw fit. Some used the completion of questionnaires on work values and leadership styles as a catalyst for a subsequent discussion of the impact of culture on organisational behaviour. Others, after debriefing the students about the dual purpose of the project, engaged students in a discussion about the role of language. However, those collaborators who did not teach classes specifically related to these topics, or who collected data in one of their colleagues' classes, simply requested the students' collaboration in an international research project. Incentives used varied from none to sweets and a prize draw for an electronic gadget.

Distribution of the questionnaires also varied slightly. In some class settings it was easy to separate languages in the same class as there were two separate aisles in the lecture theatres, in others questionnaires were alternated by row. In countries where problems were expected with distributing questionnaires in different languages in the same class, day and evening classes or repeat streams were used to separate languages. We tested for differences in demographics (see above), and hence we could be confident that these local adjustments had not impacted on the soundness of the research design. However, local variation can go too far and needs to be closely monitored. The Japanese collaborator had carefully instructed her colleagues in whose classes the data were collected on the aim of the study. Unfortunately, they then proceeded to ask students in which language they wanted to complete the questionnaire, hence completely obstructing the randomisation process. Because of small group sizes, the French collaborator was forced to distribute the different languages in groups with very different types of students (e.g., Masters students vs. Research students). Hence, we were unable to use the data for these two countries for the test of the language effect. Close monitoring of adherence to research designs becomes more difficult as the number of countries increases. Therefore, although it is tempting to continue to add more countries to a data set if collaborators volunteer to be involved, it might be better to focus on a more limited number of countries that are purposefully chosen according to pre-specified theoretical criteria.

In project 2 data were collected locally in each country (with the exception of France and the Nordic countries, for which questionnaires were mailed from the UK). One of the research team members travelled to each of the Asian countries and collected data as a visiting scholar at prestigious local institutions. This team member also collected data in Germany and the UK whilst employed at institutions there. Data for Australia and Spain were collected by other members of the research team working in these respective countries. In each country, we made slight adjustments to the data collection procedures and incentives used. In most countries our questionnaires included pictures of screenbeans encouraging the respondent to carry on with the questionnaire. In countries such as Germany or France, however, we refrained from doing so as this might have been perceived as unfitting for an academic study. In Australia and China we promised donations to a charity for each questionnaire that was returned to us, the Cancer foundation in Australia and a charity relating to a recent major earthquake in Sichuan province in China. In Korea, we employed a local survey firm who telephoned respondents to ensure the higher response rate necessary to compensate for the smaller population of foreign-owned firms in Korea. Given that one of the team members travelled to Japan, China and Korea to administer the survey and was affiliated as a visiting scholar at leading local business schools, he was able to send out the invitation letter on letter-headed paper of these institutions, thus "localising" the survey to some

extent. In both Japan and France, letters of endorsement were provided by professors at prestigious institutions. In the UK we included tea bags, suggesting in our invitation letter that receivers should take a break, have a cup of tea and fill out the questionnaire in the meantime (see also Harzing, 1997). In Spain, another team member checked the company address list against the Alumni database of his home institution, a highly prestigious business school with a large Alumni network in Spain, to identify the names of the respective HR directors. This helped to develop personalised letters and increase response rates. Further, in Australia, Germany and the UK letters were sent out from prestigious home institutions which should have given the survey additional weight. In Germany, the researcher – who had just moved from the UK to Germany – used the opportunity to introduce himself to the local business community and promised a business report. No specific incentives were offered in the Nordic countries because based on prior experience (Harzing, 1997) we expected relatively high response rates in these countries. It is quite likely that the lack of incentives, the fact the questionnaire was not localised, and the mailing of the questionnaire from the UK led to mere average response rates in the Nordic countries (see below). In terms of personalising our correspondence with potential respondents, we signed each letter by hand, even though it meant signing more than 10,000 letters.

Illustration 7: Differences in response rates between countries

Our overall response rate in project 2 was 13.8%, but this includes a very high response rate for Korea (47%) where we used telephone surveying through a survey company. Excluding Korea our response rate was 9.6%. Double-digit response rates were achieved for Japan (10.4%), Germany (11.1%), Nordic countries (11.3%), Australia/NZ (12.7%) and Spain (15.4%). China (4.0%), the UK (5.2%) and France (6.6%) had much lower response rates.

The low response rate in China is likely to have been caused by China being the only country in which we did not send out paper questionnaires, as well as the fact that our web survey was blocked in the initial mailing. In addition, the lack of experience with academic research in China might have negatively influenced response rates. Furthermore, culture might also have played a role. In universalistic countries, academic questionnaires tend to be filled out more often, as this assists in the generation of a "greater good", i.e. knowledge gain in academia. By contrast, in particularistic countries, and China is a case in point, a favour is more likely to be done only to people one has a direct connection with. In the UK, low response rates might have been caused by the fact that data were collected in the height of the first wave of the Global Financial Crisis, which no doubt led HR managers to have other priorities than completing questionnaires. In multi-country studies France is typically one of the countries with the lowest response rates (see Harzing, 1997). Not sending questionnaires locally (questionnaires were mailed from the UK and Australia) might also have negatively influenced response rates in France, although we did include a recommendation letter from France's most prestigious business school.

Given that respondents were directly recruited through their participation in non-degree granting executive education programs, the response rate in project 3 was substantially higher and reached 34.6% among the participating managers. Again, however, manager response rates by country differed substantially, from very high rates in Peru (91.1%), Norway (83.3%), Romania (78.5%) and Greece (76.7%) to low rates in Russia (20.3%), Brazil (16.8%), Ireland (11.2%) and Colombia (10%). However, these

differences are more likely to be due to factors such as the local institution's reputation or the number of participants per executive education program than to inherent country differences in survey response rates.

Summary of suggestions

Concerning the international survey and data collection process we recommend that researchers pay careful attention to and explicitly cater for possible cultural differences in the perception of survey design, survey administration and incentives offered to participants. Involving local collaborators and/or pilot-testing the survey can help researchers to do so. In addition, personalising the invitation and reminder letters, and obtaining sponsorship from local institutions can help increase response rates.

5 Data analysis

In international and cross-cultural research, the effect of cultural differences has to be explicitly taken into account in order to draw meaningful inferences from the survey results. In this regard, several statistical approaches have been developed to test for and establish cross-cultural equivalence. A first set of techniques are based on item response theory which examines statistical relationships between item responses and the latent attributes that are reflected by combinations of specific items. If these statistical relationships and thus item response distributions reveal similar patterns for constructs measured in different languages, it is assumed that construct equivalence is possible (Peng et al., 1991). In a different vein, Riordan and Vandenberg (1994) apply a covariance structure analytic procedure to test the stability and transferability of self-report measures in cross-cultural research. Similarly, Mullen (1995) applies Multiple Group LISREL and Optimal Scaling techniques to the diagnosis of cross-cultural equivalence. However, a main drawback inherent in these methods is the need to have equally-sized groups in order to model comparisons which may be difficult to achieve when multiple cultural groups are considered. As mentioned earlier, local collaborators and even other local academics volunteering to peer-review the results can serve as an important source for interpreting the findings within the scope of the local cultural and institutional context.

A rather vexing problem in cross-national research is the issue of response style differences across countries. Studies of attitudes across countries have generally relied on a comparison of aggregated mean scores to Likert-scale questions. This presupposes that when people complete a questionnaire, their answers are only based on the substantive meaning of the items to which they respond (Baumgartner and Steenkamp, 2001). However, people's responses are also influenced by their response style. 'Response style' refers to a respondent's tendency to respond systematically to questionnaire items regardless of their content (Baumgartner and Steenkamp, 2001). The most commonly cited examples of response styles are acquiescence (ARS) or disacquiescence (DRS); that is, the tendency to agree or disagree with an item regardless of the content, and extreme response styles (ERS) versus middle response styles (MRS); that is, the tendency to use the extreme or middle response categories on ratings scales.

Prior research has shown that there are differences in response styles across countries, especially for attitudinal questions such as cultural norms and values (see e.g., Harzing, 2006; Smith 2004). Smith (2004) investigated acquiescence bias and used 27 cultural

dimensions as explanatory factors on this response bias. He found that two of the Hofstede (1980) dimensions, Collectivism and Power Distance, and two of the Globe dimensions, Uncertainty Avoidance "should be" and Family Collectivism "as is", were consistently and strongly positively related to ARS across a wide range of cross-cultural studies. These findings were subsequently confirmed by Harzing (2006) in a study with student respondents across 26 countries.

Although less information is available about middle and extreme responses styles, Harzing (2006) found middle response styles to be more frequent in collectivistic countries, whereas country level extraversion was related to extreme response styles. The same study also found extreme responses to be more likely when a respondent is responding in his or her native language, whereas middle responses were more likely when English language questionnaires were used. Harzing et al. (2012) found that Asian respondents showed higher MRS than Western respondents. When scale anchors referred to naturally opposing and mutually exclusive constructs, respondents showed more ERS than when they referred to level or degree of a construct. Knowledge of cross-national differences resulted in higher ERS on behavioural questions.

These results show that researchers should always test whether response styles are present before further analysing their results. What might be construed as a higher mean score about the topic in question might simply be an acquiescence bias. Alarm bells should certainly start ringing when one country group has consistently higher mean scores for any of a set of unrelated constructs. There are various ways to address response bias in cross-national studies; the most common of which is standardisation of responses (see Fischer, 2004). Other solutions all relate to initial questionnaire design. A use of a mixture of positive and negative statements will mitigate both acquiescence and disacquiescence. Likert scales with a larger number of scale points and the use of ranking have also been shown to reduce both response and language bias (Harzing et al., 2009) as have scale anchors that refer to mutually exclusive constructs, rather than to level of agreement (Harzing et al., 2012).

Illustration 8: Standardisation to address response style differences for "subjective" and "objective" questions

In project 2, our questions were not strictly speaking attitudinal questions as they dealt with company practices. However, as we asked key informants (HR managers) to respond on behalf of the company there might be a perceptual element to the result. We had already decided to use 7-point scales in our study as they have been shown to perform better than 5-point scales in terms of attenuating response style effects (Harzing et al., 2009). However, when analysing the results, we noticed that consistently higher scores appeared to be given for nearly all questions by our Chinese, Korean and Spanish respondents.

We therefore averaged the scores across all 7-point Likert scale questions in our survey. Whilst the average score for the countries/regions was close to the theoretical mean score (4 on a 7 point scale), China, Korea, and Spain had significantly higher average scores (t=6.422, p=0.000). Further, these three countries are significantly different from the other six countries in our survey on the very four cultural dimensions that Smith (2004) and Harzing (2006) had identified as the main determinants of acquiescent response bias. Hence our results indicate that even for non-attitudinal questions response style differences between countries might distort comparisons of mean scores between countries for the constructs under investigation.

Some authors have argued that cross-cultural differences in response bias should be seen as differences in communication styles across countries rather than bias to be controlled for (see e.g., Smith, 2004). Whilst this interpretation might be appropriate for studies looking at individual cultural norms and values or other attitudinal responses, we do not think this reasoning is valid for response style differences with regard to "objective" company level constructs such as knowledge transfer, performance and autonomy. To address the acquiescent response bias, we therefore standardised all constructs by subtracting the respondent's average mean score from the construct's raw score. This is called within-subject standardisation (Fischer, 2004). We did not find significant differences between countries on the average variance across questions and hence extreme or middle response bias did not appear to be a problem.

Summary of suggestions

In the data analysis and publication stage of international survey research we recommend scholars to conduct a set of measurement equivalence tests as part of the preliminary analyses and clearly explain their use in the later write-up of the article. Additionally, researchers should test whether response styles are present and deal with these biases both (1) *a priori* by combining positive and negative item statements, using a larger number of scale points, employing ranking instead of rating, and using scale anchors that reflect mutually exclusive constructs, and (2) *post hoc* through the standardisation of responses.

6 Publication of results

The final stage in the research process involves the publication of the research findings. An important part of international collaborative research is to establish a clear publication strategy and determine co-authorship at the outset to avoid disappointments (see Teagarden et al., 2005). This is particularly relevant if the research team involves a multitude of scholars that do not know each other personally or only very rarely meet. This also entails deciding on possible target journals early on in the project. For example, if an international research team chooses to publish in US journals, then this choice is likely to influence discussions about specific topics and methods employed (Peterson, 2001). It is also important to be aware of different power relations within the research team (Easterby-Smith and Malina, 1999). Although every project will require leadership by one or a few researchers, these principal researchers are in a position of power because they are often the only ones in control of the full data set and the aggregate data analyses. In contrast, local collaborators hold expert power through their control over and understanding of local data, which allows them to also publish independently of the principal researcher. It is therefore essential to define co-authorship rules for all publications resulting from a specific collaborative project.

Many international collaborative research efforts involve the collection of large databases with the aim of publishing multiple papers. Obviously, one does not engage in a 3–5 year data collection process in 10–25 countries to just write up a single paper. Even so, many top journals are becoming increasingly strict about publishing multiple papers from the same data set. To avoid overlap between papers based on the same data set, scholars have suggested clearly separating the variables under study in each of these papers and adopting different theoretical angles (Kirkman and Chen, 2011).

Finally, when writing up the results for publication, it is important to explicitly discuss and convince potential reviewers and editors why the research project warrants an international rather than a domestic design, especially in the case of targeting general management and organisation studies outlets. Of course, a unique research contribution should be established prior to starting data collection, but an explicit justification of conducting international survey research along with a discussion of the distinct methodological issues provides a stronger rationale that reviewers and editors, and in fact readers at large may buy into.

Illustration 9: Publication strategies in multi-country projects to manage collaborators' expectations and optimise research output

Project 1 involved a large number of country collaborators. For most of these collaborators it was more important to have publications quickly than to have publications in the highest level journals, so we produced lots of conference papers, not all of which resulted in journals publications. We also made specific agreements about co-authorships from the start. One cannot expect country collaborators to put in a significant effort in data collection if they are only recognised in the acknowledgements. For this project, the agreement was thus that all country collaborators would be co-authors on the first paper and any conference papers leading up to it, without having to do any actual writing. After that, all country collaborators were free to suggest papers and in fact two of the country collaborators took the lead to write up a number of content related papers.

More generally, it is important to realise that managing large research teams requires significant dedication and patience, as well as a delicate balance between democracy and clear direction (see Teagarden et al., 2005 for an excellent overview of the issues in each of the stages of the life cycle of multinational research teams). Thousands of emails were exchanged throughout the life cycle of project 1 and considerable cross cultural empathy was required from all in dealing with the inevitable differences with regard to time management and email communication. Throughout the project, the coordinator also provided regular status reports to keep the collaborators informed of any progress or just to alert them to problems encountered on the way. For important decisions, decision-making was expressly democratic, even if many collaborators did not take the opportunity to have a say. To facilitate social cohesion, all collaborators were asked to write up a short personal story about themselves that was shared amongst the whole team. Some collaborators proceeded to work together on other projects, including project 3 described in this article.

In project 2, collaboration was much more straightforward in that all three team members had worked together before and shared similar cultural backgrounds and work ethic. Partly as a result of this, we were also able to follow a different publication strategy. In this project, we purposefully mapped out all possible papers from the project in an iterative process taking nearly a year. As a consequence we were able to: ensure that the papers have no overlap in terms of variables or theory, create a very clear timeline, clearly distinguish authorship responsibilities, as well as have a coherent strategy in terms of conference presentations and journal outlets. Obviously, this type of structured approach only works if collaborators share similar goals and work practices.

In many ways, project 3 was similar to project 1 in that it involved many collaborators at an early career stage for whom timely research results were very important. The project resulted in several conference presentations that were co-authored

by all collaborators, listed in alphabetical order after the core team of researchers responsible for analysing the data and crafting the paper. A similar agreement was reached for the first journal paper reporting the main quantitative results of the study. Any additional papers could then be proposed and written up by a sub-team of country collaborators. In addition to journal articles, the project also resulted in a joint book about trust in different cultures that would report the quantitative and, importantly, the rich qualitative results per country. Here, local collaborators were responsible for writing up their respective country chapters in exchange for co-authorship in the book. A standardised structure for each chapter was decided on and a chapter template for one country was shared with the whole team before the remaining chapters were written up.

Summary of suggestions

To publish international survey research it is important to clearly determine a publication strategy and rules of co-authorship for all collaborators. In many regards, managing an international team of researchers can be likened to managing a global virtual team that requires multiple points of contact to increase mutual trust and collaboration. It is also important at the outset of the project to be conscious about how multiple papers from the same data set can be crafted without risking too much overlap in variables and theoretical perspectives. Finally, scholars should explicitly highlight in the write-up of their research findings in which way the study design is distinct from a domestic research context.

7 Conclusion

Conducting meaningful international empirical research is prone to additional difficulties and complexities and can easily discourage researchers from initiating cross-cultural inquiry in the first place. More specifically, international researchers who collect primary data, either through questionnaires, interviews or other means, are invariably confronted with language barriers, cultural barriers, geographical distance and the liability of foreignness, which all result in higher monetary costs and a more significant time investment. This can easily lead to a lower research output compared with that of researchers who either limit themselves to the familiar domestic context, or employ secondary data and are therefore not confronted with these obstacles. If it comes to recruitment or promotion decisions, these systemic disadvantages for international business researchers are often not sufficiently taken into account by the employing institution. In addition, given that international survey research is frequently associated with high monetary costs and a very significant time investment, it only makes sense to produce several papers from one project. However, many top journals are hesitant to accept papers coming from the same data set as previously published papers. This practice could easily lead to researchers to focus on small-scale, incremental and piecemeal projects instead, which would harm the advancement of knowledge.

On a related note, we would like to encourage researchers to be more open about their research process. Our research publications tend to be mainly about our results. Even sections on methodology are formalised and standardised and do not typically illustrate the messiness of research. Furthermore, if international business scholars attempt to produce more than one paper from a particular data set, they see themselves forced to be “original” for each of these contributions, even in the methodology section,

in order to avoid self-plagiarism. This is regrettable, given that sections on data collection, sample, and questionnaire development could de facto be very similar between contributions.

Furthermore, typically only successful research projects, i.e. those that produced tangible and significant results get published, whereas unsuccessful ones do not leave any trace except as bad memories for the researchers involved. However, even unsuccessful projects might carry important lessons in terms of both results and research methods that could be worth reporting. As one of our Masters students said: If you see a good example, you read it and think: "Well, I would have done exactly the same" and just go on with what you were doing; if you see a bad example, you think: "Oh dear, perhaps I would have done exactly the same", and learn from it. Consequently, we would encourage journals to be more open towards the publication of research that would conventionally be considered as "failed", as we might still be able to learn useful lessons from it. This would have the added advantage of preventing scholars from adapting their hypotheses ex post to the data in order to turn their failed research project into a "successful" one. One might even envisage a journal that focuses on lessons to be learned from "failed" projects.

Given the particular challenges of international survey research we described above, many areas in the field of international management still remain largely under-researched, even though they provide ample opportunities to advance our knowledge. However, we hope that by identifying some of the key issues in international survey research and offering various solutions, we have been able to encourage and promote such future research.

Acknowledgements

Part of the research on which this article is based was supported under Australian Research Council's Discovery Projects funding scheme (project DP0555977). The views expressed herein are those of the authors and are not necessarily those of the Australian Research Council.

References

Adler, N.J., Campbell, N. and Laurent, A. (1989) 'In search of appropriate methodology: from outside the People's Republic of China looking in', *Journal of International Business Studies*, Vol. 20, No. 1, pp.61–74.

Akkermans, D., Harzing, A.W. and Witteloostuijn, A. van (2010) 'Cultural accommodation and language priming. Competitive versus cooperative behavior in a prisoner's dilemma game', *Management International Review*, Vol. 50, No. 5, pp.559–584.

Baruch, Y. (1999) 'Response rate in academic studies – a comparative analysis', *Human Relations*, Vol. 52, No. 4, pp.421–438.

Baumgartner, H. and Steenkamp, J-B.E.M. (2001) 'Response styles in marketing research: a cross-national investigation', *Journal of Marketing Research*, Vol. 38, pp.143–156.

Bello, D., Leung, K., Radebaugh, L., Tung, R., and van Witteloostuijn, A. (2009) 'From the Editors: student samples in international business research', *Journal of International Business*, Vol. 40, pp.361–364.

Brislin, R.W. (1970) 'Back-translation for cross-cultural research', *Journal of Cross-Cultural Psychology*, Vol. 1, No. 3, pp.185–216.

Brislin, R.W. (1980) 'Translation and content analysis of oral and written material', in Triandis, H.C. and Berry, J.W. (Eds): *Handbook of Cross-Cultural Psychology: Methodology*, Vol. 2, Allyn and Bacon, Boston, MA, pp.389–444.

Caprar, D. (2011) 'Foreign locals: a cautionary tale on the culture of MNC local employees', *Journal of International Business Studies*, Vol. 42, pp.608–628.

Cycyota, C.S. and Harrison, D.A. (2006) 'What (not) to expect when surveying executives: a meta-analysis of top manager response rates and techniques over time', *Organizational Research Methods*, Vol. 9, No. 2, pp.133–160.

Dillman, D.A. (2006) *Mail and Internet Surveys: The Tailored Design Method*, 2nd ed., Wiley, New York.

Easterby-Smith, M. and Malina, D. (1999) 'Cross-cultural collaborative research: toward reflexivity', *Academy of Management Review*, Vol. 42, No. 1, pp.76–86.

Fischer, R. (2004) 'Standardization to account for cross-cultural response bias: a classification of score adjustment procedures and review of research in JCCP', *Journal of Cross-Cultural Psychology*, Vol. 35, No. 3, pp.263–282.

Fricker Jr., R.D. and Schonlau, M. (2002) 'Advantages and disadvantages of internet research surveys: evidence from the literature', *Field Methods*, Vol. 14, No. 4, pp.347–367.

Harpaz, I. (2003) 'The essence of performing meaningful comparative international survey research', in Punnett, B.J. and Shenkar, O. (Eds): *Handbook for International Management Research*, 2nd ed., University of Michigan Press, Ann Arbor, MI, pp.17–48.

Harzing, A.W. (1997) 'Response rates in international mail surveys: results of a 22 country study', *International Business Review*, Vol. 6, No. 6, pp.641–665.

Harzing, A.W. (1999) *Managing the Multinationals: An International Study of Control Mechanisms*, Edward Elgar, Cheltenham.

Harzing, A.W. (2000) 'Cross-national industrial mail surveys', *Industrial Marketing Management*, Vol. 29, No. 3, pp.243–254.

Harzing, A.W. (2006) 'Response styles in cross-national survey research: a 26-country study', *International Journal of Cross-Cultural Management*, Vol. 6, No. 2, pp.243–266.

Harzing, A.W. et al. (2005) 'Does the use of English-language questionnaires in cross-national research obscure national differences', *International Journal of Cross-Cultural Management*, Vol. 5, No. 2, pp.213–224.

Harzing, A.W. et al. (2009) 'Rating versus ranking: what is the best way to reduce response and language bias in cross-national research?', *International Business Review*, Vol. 18, No. 4, pp.417–432.

Harzing, A.W., Brown, M., Köster, K. and Zhao, S. (2012) 'Response style differences in cross-national research: dispositional and situational determinants', *Management International Review*, Vol. 52, No. 3, pp.341–363.

Hewson, C., Yule, P., Laurent, D. and Vogel, C. (2003) *Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences*, Sage, London.

Hinkin, T.R. (1995) 'A review of scale development practices in the study of organizations', *Journal of Management*, Vol. 21, No. 5, pp.967–988.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-Related Values*, Sage, Beverly Hills, CA.

Hofstede, G. (1998) 'Think locally, act globally: cultural constraints in personnel management', *Management International Review*, Vol. 38, No. 2, pp.7–26.

Hult, G.T.M., Ketchen, D.J., Jr., Griffith, D.A., Finnegan, C.A., Gonzalez-Padron, T., Harmancioglu, N., et al. (2008) 'Data equivalence in cross-cultural international business research: assessment and guidelines', *Journal of International Business Studies*, Vol. 39, No. 6, pp.1027–1044.

Kirkman, B.L. and Chen, G. (2011) 'Maximizing your data or data slicing: recommendations for managing multiple submissions from the same data set', *Management and Organization Review*, Vol. 7, No. 3, pp.433–446.

Marschan-Piekkari, R. and Welch, C. (Eds) (2004) *Handbook of Qualitative Research Methods for International Business*, Edward Elgar, Cheltenham.

McGrath, J.E. (1982) 'Dilemmatics: the study of research choices and dilemmas', in McGrath, J.E., Martin, J. and Kulka, R.A. (Eds): *Judgment Calls in Research*, Sage, Beverly Hills, CA.

Mullen, M.R. (1995) 'Diagnosing measurement equivalence in cross-national research', *Journal of International Business Studies*, Vol. 26, No. 3, pp.573–596.

Olivas-Luján, M.R., Harzing, A-W. and McCoy, S. (2004) 'September 11, 2001: two quasi-experiments on the influence of threats on cultural values and cosmopolitan-ism', *International Journal of Cross-Cultural Management*, Vol. 4, No. 2, pp.211–228.

Peng, T.K., Peterson, M.F. and Shyi, Y-P. (1991) 'Quantitative methods in cross-national management research: trends and equivalence issues', *Journal of Organizational Behavior*, Vol. 12, No. 2, pp.87–107.

Peterson, M.F. (2001) 'International collaboration in organizational behavior research', *Journal of Organizational Behavior*, Vol. 22, No. 1, pp.59–81.

Riordan, C.M. and Vandenberg, R.J. (1994) 'A central question in cross-cultural research: do employees of different cultures interpret work-related measures in an equivalent measure?', *Journal of Management*, Vol. 20, No. 3, pp.643–671.

Sekaran, U. (1983) 'Methodological and theoretical issues and advancements in cross-cultural research', *Journal of International Business Studies*, Vol. 14, No. 2, pp.61–73.

Sheehan, K. (2001) 'E-mail survey response rates: a review', *Journal of Computer-Mediated Communication*, Vol. 6, No. 2.

Smith, P.B. (2004) 'Acquiescent response bias as an aspect of cultural communication style', *Journal of Cross-Cultural Psychology*, Vol. 35, No. 1, pp.50–61.

Teagarden, M.B., Drost, E.A. and Von Glinow, M.A. (2005) 'The life cycle of academic international research teams: just when you thought "virtual" teams were all the rage … here come the AIRTS!', *Advances in International Comparative Management*, Vol. 18, pp.303–336.

Tomaskovich-Devey, D., Leiter, J. and Thompson, S. (1994) 'Organizational survey non-response', *Administrative Science Quarterly*, Vol. 39, No. 3, pp.439–457.

Tse, A. and Ching, R. (1994) 'A comparison of the effectiveness of mail and facsimile as survey media on response rate, speed and quality', *Journal of the Market Research Society*, Vol. 36, No. 4, pp.349–355.

Tsui, A.S., Nifadkar, S.S. and Ou, A.Y. (2007) 'Cross-national, cross-cultural organizational behavior research: advances, gaps, and recommendations', *Journal of Management*, Vol. 33, No. 3, pp.426–478.