
Construct Equivalence in International Business Research: The First and the Last of It

**David N. McArthur*
Utah Valley State College

Construct equivalence is a necessary condition for establishing validity in comparative cross-cultural studies. This paper explores the relationship between equivalence and validity as presently understood in international business literature from management and marketing writers' work, and proposes ways that accounting and finance researchers can become alert to and deal with threats to equivalence and validity in their international business research.

Key Words: Construct equivalence, Validity, International research

Introduction

Constructs are the mental models of phenomena through which the world is understood. A construct differs from a concept in that a researcher develops a construct in order to study questions about dimensions, time, and content of links between the present construct and the environment and other constructs. Published research in international business (IB) often studies phenomena through the author's constructs and analyses by comparing data from different geo-political jurisdictions (nation-states) which too frequently act as proxies for cultures. IB researchers in cross-cultural comparative projects may expect subjects to have the same understanding of business-related constructs when in fact different purposes exist in different cultures (e.g., strategic planning in Japan and the USA, management delegation in Canada and Turkey, or bankruptcy in Germany or the USA). Differences may include such things as different meanings, the temporal nature of links to other constructs, or content between links.

It is elemental that the construct or mental model of a phenomenon influences the research design (Wallace, 1971). Design, in turn, influences the way research is undertaken and results are obtained. Results influence analysis and its interpretation, inferences, and conclusions drawn (as do the research design and the construct or the phenomenon). Results should reflect

not only solid measurement and analysis but should stand up to replication and practical use. Cross-cultural research assumes there are inter-cultural differences in the domain of the construct or model of interest. If the construct is not equivalent in both (or all) of the cultures being compared, the resulting data will be confounded (Jacoby, 1978) in part because inter-cultural differences will not be the only source of variance. Thus, construct equivalence is a necessary condition for construct validity in cross-cultural scientific studies. This was brought to the fore in recent articles by Singh (1995) and Mullen (1995) which re-examined data sets of earlier studies on job satisfaction and organizational commitment (Lincoln & Kalleberg, 1985) and role conflict, role ambiguity, job performance, job satisfaction, and organizational commitment (Dubinski, Michaels, Kotabe, Lim, & Moon, 1992). In both cases, re-examination of the data revealed that the earlier researchers had suffered from construct inequivalence not corrected by simple back translation and had made both type I (mistakenly thinking there was a significant difference across countries, when there was not) and type II errors (concluding there was no mediating relationship between culture and the dependent-independent variables when there was).

Much comparative research assumes construct equivalence (Adler, 1983, Boyacigiller & Adler,

1991, Singh, 1995). Some researchers may assume that the familiar interpretation and meaning of concepts in accounting, finance, marketing, organization science, and management are *givens* that can be expected of respondents in other countries (or at least in the country where they will gather data).

This intellectual bias of convenience is obvious to the critical reader as shown by Boyacigiller and Adler (1991, p. 266) who, after reviewing the international research literature in management, observed, "The scope and primary orientation of most theories are American; however, they are presented as if they were universally applicable."

Other research (e.g., Kotabe, 1992) explains the systems being compared and then uses secondary data to evaluate the systems for differences. A problem here is that one person's secondary data was first another person's primary data. Those who gather primary data do so with their construct in mind, in their design. Those who interpret that as secondary data must assume equivalence. Bhagat and McQuaid (1982, p. 680) insist that, at least for primary data collection using questionnaire methods, construct equivalence should be well established.

Construct Equivalence

"To compare two phenomena, they must share some feature in common; and to compare them to some advantage, they should usually differ on some feature ..." (Berry, 1980).

At a superficial level, construct equivalence in cross-cultural research is a condition where a foreign respondent understands the data-gathering question as the researcher intends it to be understood. The ensuing answer, if generalizable, confirms the comparative differences proposed in the researchers' model. The manifestation of this equivalence lies in the construct validity of the instrument (issues of internal and face validity of the instrument are dealt with in Brinberg and McGrath, 1985; Cook and Campbell, 1976). Construct equivalence is a vital condition for primary cross-cultural data to be comparable and for results to be interpretable (Douglas & Craig, 1983).

Construct equivalence and validity are of concern to many fields of research whether or not the domain is

cross-cultural. Articles such as those by Venkatraman and Grant (1986) in organizational strategy, Bagozzi, Yi, and Phillips (1991) in organization science, and Nasif, Al-Daeaj, Ebrahim, and Thibodeaux (1991) in cross-cultural management offer critical advice to researchers in several literatures. Early consideration of the ways the constructs at the heart of a research project might differ in meaning in foreign environments or differ in linkage with other constructs will go a long way toward establishing construct equivalence and validity. For research already underway, a worthy intermediate goal might be the lesser target of measurement validity.

Cross-cultural research literature outlines four dimensions along which concepts exhibit equivalence (Frijda & Jahoda, 1966; Berry, 1969, Sears, 1951, Douglas & Craig, 1983). They include functional, conceptual, categorical, and metric equivalence. Writers in the 1980s (Douglas & Craig, 1983; Berry, 1980; and Brislin, Lonner, & Thorndike, 1980) began to agree on two basic dimensions (functional and conceptual) but treat the remaining dimensions differently.

Functional equivalence (Frijda & Jahoda, 1966) exists when behaviors are observable in two cultures that relate to functionally similar problems. Berry (1969) stressed functional equivalence when he pointed out that concepts or constructs may not serve the same role or function in different cultures. Douglas and Craig (1983 p. 137) illustrate the problem presented by functional inequivalence by pointing out the different uses of a bicycle in the Netherlands where it frequently serves as basic transportation and in the USA where its main use is for recreation. As market researchers begin to examine the relevant market segments and competing product sets for home country products, they have to rethink the membership and structure of each because of different uses and the meanings those uses have in research on decision making.

Sekaran (1983) wrote of the need for researchers to ensure that their research reflects functional equivalence (p. 62). She questioned its importance however, since IB "researchers usually study comparable work settings in different cultures. Thus, this issue (functional and construct equivalence) may not be critical." Some researchers feel that this

problem is not as bad in accounting or finance as it is in marketing but these disciplines may have research conclusions and received wisdom that are wrong.

Conceptual equivalence (Sears, 1951) is not viewed as a function of the roles filled by the constructs used in the research design but, rather, as the meanings of concepts, constructs, objects, or phenomena in a particular culture. Conceptual meaning must be equivalent before comparison of cross-cultural data is possible, therefore researchers must find the meanings within the cognitive systems of the cultures being studied. Often these meanings are found in probing the links between a construct and others with which it naturally occurs in the foreign setting. These cognitive systems, with constructs and ties to other constructs and to the factual environment, are the major structure on which people and organizations make sense of or categorize experience. Equivalence exists if there is an underlying similarity in the dimensions along which people group phenomena in like categories.

Conceptual equivalence may be problematic in international research in all disciplines. Management researchers must reevaluate expressions of personality in different cultures, marketers have to “discover” local ways cultures express social interaction, and financial researchers have nonequivalent concepts at work when they compare secondary data on bankruptcies or linkages between firm performance and stock price between Germany, China, the USA or other countries.

Category equivalence (Douglas & Craig, 1983) is a specialized type of conceptual equivalence detailing the ways in which subjects (e.g., consumers) categorize phenomena (e.g., brands or products). Category equivalence emphasizes the rules, logic, or grammar by which persons from a culture “cut up the pie of experience” (Triandis, 1972, p. 10). Persons within a culture tend to “cut” or categorize a phenomena similarly. Groups from different cultures may have significant differences in their patterns of categorization.

Douglas and Craig (1983) refer to metric equivalence as a reflection of the measurement instrument and not of the underlying concept, while Berry (1980) includes it as a component part of conceptual equivalence. Metric equivalence occurs when the

study data exhibits similarities of structure within cultures close enough to allow researchers to reasonably assign the majority of the remaining variance to inter-cultural differences. Metric equivalence rests on the assumption that respondents from other cultures being compared understand the scale devices and respond to them in similar fashions. Scales or scoring procedures in quantifying measures differ across cultures. Behavioral scientists in North America are used to using five- or seven-point scales (such as Likert items) but other countries or regions may be used to assessing items on ten- or twenty-point scales. Additionally, metric equivalence includes answers to questions of whether the quantitative scale used, be it a behaviorally anchored rating scale comparing management effectiveness in South Korea and the US or measures of firm distress that are used to predict bankruptcy between US and Japanese publicly traded firms, has similar meanings as it quantifies some measurable phenomenon in each country (Drasgow, 1987).

Emic and Etic Approaches to Conceptualization

Another problem in construct equivalence is the pervasive use by cross-cultural researchers of emic, etic, and pseudo-etic (or imposed etic) approaches (Pike, 1967; Triandis, 1972; Triandis and Marín, 1983; Brislin, Lonner, & Thorndike, 1980; Berry, 1969).

“Emic” and “etic” are terms taken by Pike (1967) from linguistics where *phonemics* denotes sounds used only within a particular linguistic system and *phonetics* are sounds used universally. Similarly, critics of cross-cultural research use “an emic approach” to mean a research approach and design that studies behavior within a culture. An “etic approach” connotes an approach and design that uses external, culture-free, universal terms to study behavior (Bhagat & McQuaid, 1982; Triandis & Marín, 1983).

The problem of inappropriate use of etic, emic, and pseudo-etic approaches may arise because unknowing researchers find they can use recognizable, observable behaviors – such as a firm’s declaration of bankruptcy in North America – to operationalized hypothetical constructs in the new setting – as they do in collecting secondary data on firm bankruptcies in Japan or Germany. In defaulting

to intuitive, familiar, or pre-validated indicators from their home culture rather than searching for either pan-cultural “universals” (Osgood, May, & Miron, 1975) or true “emic” indicators from the local culture, researchers fail to gain equivalence and thereby lose validity.

Triandis and Marín (1983) point out a critical error in how researchers approach cross-cultural research by focusing on errors in procedural assumptions that they call pseudo-etic or imposed etic approaches. A researcher using a pseudo-etic approach utilizes an instrument developed in one culture (e.g., accounting’s construct of ‘fair presentation’ under US Generally Accepted Accounting Practices or GAAP) on subjects in other cultures without construct validation or without including items specific to the target culture. This focus on procedural errors (e.g., the use of pseudo-etic instruments) deflects attention from the real problem of construct inequivalence. Many cross-cultural researchers are unaware of the need to examine their own epistemological assumptions and the effect of those assumptions on the research they perform.

Construct Inequivalence in Research

As early as 1983, Adler (1983) had surveyed the mainstream management literature focusing on cross-cultural literature and concluded that journal editors, scholars, professors, and practitioners needed to question whether emic research products developed from US data would hold in other cultures. Peng, Peterson, and Shyi (1991) dovetailed with Adler’s study of cross-cultural research trends, finding that the rate of research picked up slightly by late 1980s. The work on improving validity from the various angles of construct equivalence has cross-cultural comparativists in marketing moving from research designs using confirmatory factor analysis and classical measurement theory toward designs that use generalizability theory to assess measurement equivalence (Sharma & Weathers, 2003; Durvasula, Netemeyer, Andrews, & Lysonski, 2006). Organizational behavior and management scholars (Singh, 1995; Mullen, 1995; and Lytle, Brett, Barsness, Tinsley, & Janssens, 1995) have begun evaluating multiple-group structural equation modeling against optimal scaling, regression, and item response theory practices. Accounting ethics researchers (Cohen, Pant, & Sharp, 1993, 1996) also

contribute to the academic discourse. However, not all areas of international business research may be current on the level of rigor and research craft needed for rigorous cross-cultural or cross-national research.

A recent repartee between two perspectives on accounting standards and their setting (Flower, 1997; Cairns, 1997; Flower, 1998) in which the first author contended that there was a hegemony of English-speaking accounting bodies working through an international standards-setting organization to pressure others (notably the accounting standards body in the EU) to accept Anglo-Saxon accounting standards and practices which the second author tried to refute. A reading of the texts suggests that writers investigating cross-cultural accounting phenomena are suffering at least partly from non-equivalent constructs. That there is misunderstanding of accounting constructs at the expert practitioner level is also clear. Glaum (2005) surveyed German managers in 1994 and 1997 on their attitudes and conclusions about German and US accounting practices. In the three years between the surveys, managers surveyed went from rating US accounting statements as inferior providers of information (relative to German) to superior providers of information relative to German statements. Many parts of international research in accounting and finance deal with aggregated, firm or country-level secondary data, and may seem to be immune to construct equivalence problems to researchers not trained in IB research. This quote from Dean and Clarke suggests that much of the prescriptive work in those fields struggles with either practitioner or scientific inequivalency of function or concept:

Understanding what is meant by ‘fair’ is critical in any unraveling of the accounting and legal phrase ‘true and fair view’ to describe an entity’s state of affairs, and ‘fair value’ measurement of an entity’s assets and liabilities. Whereas those phrases are ubiquitous in commercial and accounting settings, the meanings attributed to them and the reliance placed on them differ greatly worldwide—will-o’-the-wisps to standard setters... (2005)

Pike (1982) argues that context matters in translating constructs from one culture to another. Berry (1980) warned that secondary data was not immune to causing errors in inferences because of preexisting

emic, etic, and pseudo-etic research designs. Their construct's context, the relationships by which it has meaning is not preserved in formulas or data tables. While all disciplines do not depend on human perception for social science data to the same degree, the advances made in construct equivalence and validity in other fields are scientifically valuable as disciplines mature and compare phenomenon internationally.

Methods Used to Improve Research Approach before Data Collection

A number of research methods have been used to approach equivalence and enhance validity. These methods build on older recommendations to use identical measures, universals, and translations of their instruments. Recently, authors have advocated combinations of these improvements in assessing metric and model equivalency in holistic approaches (Mullen, 1995; Singh, 1995; Lytle, et al., 1995).

Identical Measures and Ipsatized Measures

The most primitive of the methods used in cross-cultural research is the use of identical measures. Some researchers assume that the language and structure of the original instrument is fixed (Candell & Hulin, 1986). Validation in one culture binds them intellectually to use it rigidly in others. For the unsophisticated, this produces results which they are unaware are spurious. The more-aware researchers understand that these results are un-interpretable because of the presence of what Brislin (1973, p. 15-16) calls "plausible rival hypotheses." Both Straus (1969) and van Raaij (1978) remind cross-cultural researchers that the use of 'identical' measures in different cultures does not insure equivalence or comparability. An improvement on the unsophisticated use of identical measures is the interpretation by the researcher of results from an identical instrument according to locally-obtained empirical standards. Although this establishes some degree of within-group equivalence, those measures cannot be compared to results from other groups; there is no assurance that the measure has any construct validity. Construct equivalence is a

necessary but not sufficient condition for construct validity.

Straus (1969) allows for an intermediate research instrument by adapting the scoring scheme to use standards developed or validated in the target culture. Thus, an identical indicator used in a second culture is ipsatized or culturally adapted by being coded or interpreted using local standards. While this may allow both within and across-culture comparisons, some data differences may be lost due to "noise."

Universals

Osgood, May, and Miron (1975) used the semantic differential method to generate and test cross-cultural universals (etic indicators) that are thought to underlie "semantic space" (p. 38-40), a concept akin to the "deep structure" cited by Werner and Campbell (1970, p. 401). This deep structure or semantic space is believed to underlie the transformations of meanings into sentences. This search for universals, so vital to valid translations, has another important use in research on the more basic levels of categorization. A notable use in IB is international consumer research examining differences in motivations, approvals, and other affective terms.

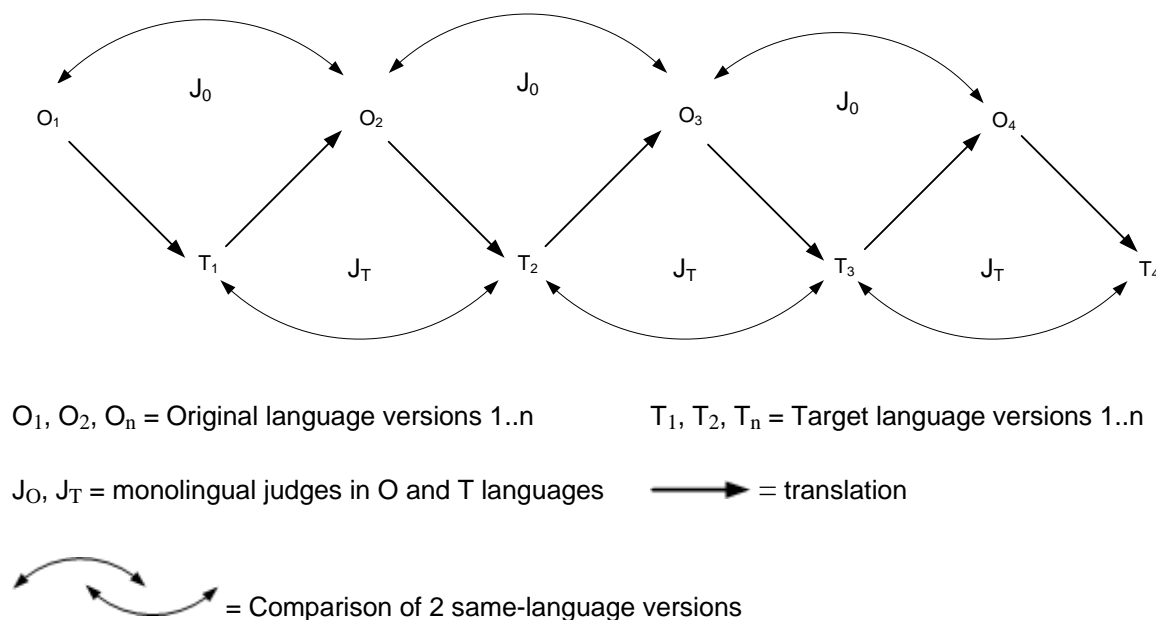
Translations, Back Translations, and Decentering

IB researchers in comparative cross-cultural studies frequently use translation of their data-collection instrument as a method to gather data from non-English speaking subjects. According to the transformational theory of language, a translator "maps" the concepts in his mind and then reformulates the sentences in the target language (Werner & Campbell, 1970). While a trained translator may understand the concept being studied, researchers need to be aware that the conceptual structures of bilinguals are different from those of monolinguals (Triandis, 1972). Thus, the translator seeks a simulacrum in the target language for a concept on the researcher's instrument which is understood in the source language. How close the translator comes, and whether or not he or she can or will express uncertainty with the translation, is problematic. Accordingly, most IB researchers use back translation to check on the validity of the first translation. It is important to remember in using back translation that slavish devotion to the source document because it has been validated elsewhere is a pseudo-etic approach.

Even better than using the back-translation for just improving the source document, Werner and Campbell (1970) advocate decentering the instrument and its translations. This practice is commonly used in cross-cultural research in many disciplines today. Decentering aims at the twin targets of loyalty of meaning and equalization of the tone of familiarity and colloquialness in each language (p. 398) – at best, decentering eliminates the distinction between source and target language, and it stresses equivalences (p. 399). Werner and Campbell advise that cross-cultural researchers first translate the source instrument (O_1) into target language (T_1), then

back-translate T_1 into the source language (O_2). An independent, monolingual judge (J_O) should then evaluate O_1 and O_2 for errors or clues to inequivalence while keeping both “targets” in mind. The researchers should then revise O_2 in line with results and again translate into the target language (T_2). Another independent, monolingual judge (J_T) should evaluate T_1 and T_2 for errors or clues to inequivalence. The researchers should keep repeating this process as long as there is significant improvement being made. Their figure (p. 415) illustrating this iterative process is reproduced in Figure 1.

Figure 1
Decentering while back-translating (adapted from Werner & Campbell, 1970)



A Composite Method for Developing Construct Equivalence

Not all comparative cross-cultural research requires translation; however, all such research does require equivalence of constructs. An alternative system of statistical correlational methods, elaborated by Przeworski and Teune (1966) and advocated by van Raaij (1978) and Brislin, Lonner, and Thorndike (1980), has been used by some to preserve the

emic-etic distinction while bringing out the meaning in hypothetical constructs.

Przeworski and Teune point out (p. 556) that identical indicators do not support equivalent inferences (see also Straus, 1969). What is needed is a set of indicators that will tap the deep meaning of a construct, its meaning, ties, and their characteristics, in each culture being compared. The first step is to devise indicators or questions which are believed by

the researchers to be etic or pan-cultural (p. 557-558, also Brislin et al., p. 27-28). Then researchers generate indicators or questions that are believed to be emic, or unique to each cultures' experience with the phenomenon. The emic or unique indicators for each culture must correlate with the etic indicators for the phenomenon. Then, as Przeworski and Teune put it,

...those indicators which are inter-correlated in a pooled, across-[culture] analysis are maintained to have identical cross-[cultural] validity with respect to a given concept and a given set of [cultures]. Those indicators which are specific to each [culture] which are correlated with the identical indicators are maintained to have equivalent cross-[cultural] validity (p. 557).

Brislin, Lonner and Thorndike (1980: 28) raise the question of how to find the list of identical indicators with which to begin the Przeworski and Teune process. They list four possible sources, namely a study of the anthropological literature of the cultures being compared in the research, first hand or exploratory contact with the culture (or, a phenomenon in a culture). They also allow that close cooperation with researchers from both cultures with which discussion of the phenomenon and development of a set of identical indicators can occur may lead to cross-cultural awareness and suitable adjustments to the research design. Lastly, they suggest discussions with expatriates or persons who have lived in both places.

Developing Construct Equivalence Holistically: Build It In and Assess It Afterward

Singh (1995, p. 610), Mullen (1995, p. 590) and Lytle, et al. (1995) build on Brislin, Lonner, and Thorndike's extension of Przeworski and Teune's work arguing for more comprehensive approaches to construct equivalence in cross-cultural research. These authors articulate and others support the view that successful cross-national research begins with

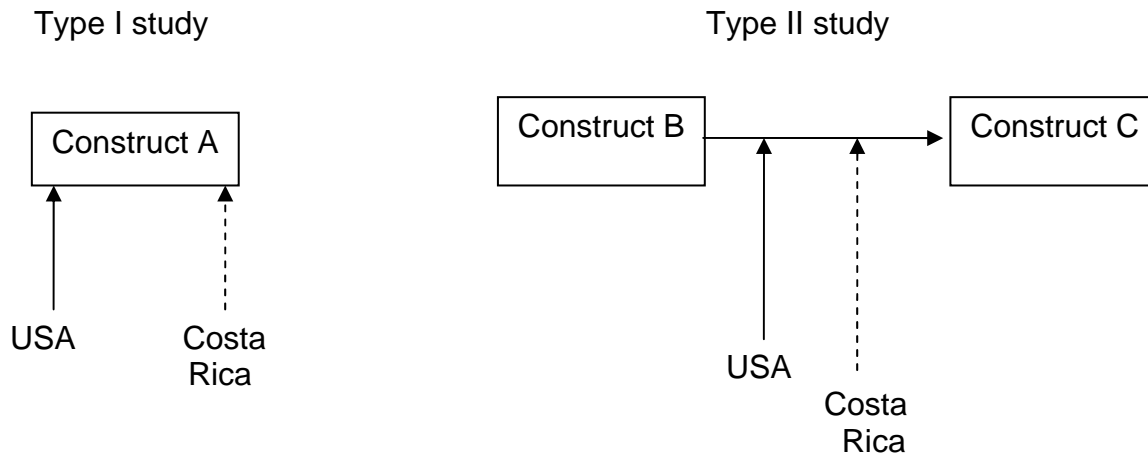
adapting traditional designs in several important areas. The traditional first step of selecting a construct (and theoretical lens) in which the phenomenon of interest is central has to be slightly skewed so that culture and its effect on the construct are included as central. This means that how cultural differences affect or moderate relationships, their causal order, or directions, are the focus of the research.

This pushes the cross-cultural researcher to review the cultural dimensions that may affect the focal construct or model reasonably (for an exhaustive review, see Lytle, et. al., 1995). One must be wary of the overlapping nature of many theoretical dimensions of culture, choosing the dimensions used in the research carefully, remembering that some dimensions are bi-polar (such as individualism – collectivism) while others are “values” (such as the need for achievement or familism).

Sample selection traditionally requires statistically based studies to follow some plan so that the sample is representative and generalizable. The cross-national researcher has the additional requirement of selecting the sample groups (the cultures or countries) so as to maximize the variation between groups on the dimension of culture that is theorized to affect the focal construct. Since cultural dimensions (and other forces) may offer competing explanations for the theorized cultural effect, during the research design one should add cultural groups and cultural dimensions (especially those that covary) to force statistical confrontation. Dimension selection for both purposes can flow deductively from the characteristics of the model or focal construct or inductively from a study of the characteristics of cultural dimensions.

Since the central focus of the cross-national research is how culture affects a focal construct or model, researchers must craft hypotheses that capture the direct or moderating effects of culture, as in Figure 2.

Figure 2
Constructing Hypotheses According to the Nature of the Relationship Posited



Using Lytle et al.'s (1995 p. 193) terminology, in a Type I study, culture causes national groups (or cultural groups) to differ in their placement *on* a construct, while in a Type II study, cultures moderate a relationship *within* a model to different degrees. As before, the researcher needs to specify *a priori* the effect the dimension of culture will have on the construct or relationship.

Culture is a latent variable, knowable only by its dimensions (Lytle, et al., 1995). These dimensions are generally measured at the individual level but generalized to the group or national level. Take the example of a writer who is investigating a phenomenon such as perception of auditor ethics measured at the individual level. If they want to make durable conclusions that are generalizable to the country, that writer must resolve differences in unit of theory (the level to which the findings are generalized), the unit of analysis (the level where the data are analyzed), and the unit of measurement (which level the data are captured at). To support this, one should match the samples between cultures or nationalities so that functional and conceptual equivalence is addressed. Additionally, within the culture, nationality, or group, the sample needs to be representative – otherwise you are drawing conclusions for Salvadoran, university educated, middle-aged, male managers and not for Salvadorans generally.

IB research standards now bring the adaptations for cross-cultural research full circle, finishing by assessing construct equivalence rigorously. Several

statistical construct equivalence assessment methods have emerged since the correlational methods developed by Przeworski and Teune (1966). Item response theory (IRT), drawing from linguistics and translation, used very large samples to create parameters by which culture groups could be statistically evaluated on their response to instruments and between group differences inferred (Hulin, 1987; Hulin, Drasgow, & Komor, 1982). Limited by its sample size requirements, the high findings of non-equivalence, and the method's inability to include local, emic items that *thicken* the between-culture differences (Peng, Peterson, & Shyi, 1991; Lytle, et al., 1995) it has been superseded. Optimal scaling (DeVera, 1986) of the data was recommended by Mullen (1995) because its rescaled values reveal the within-group structure of the underlying metrics. However, there are no statistical tests to determine if the optimal scaled values are significant and so construct equivalence can only be hoped for, not inferred.

Currently the gold standard for assessing construct equivalence in international business research after-the-fact is multiple-group structural equation modeling (SEM) in which culture is treated as a latent variable and understood by its dimensions (paths, directions, and path coefficients). In multiple-group SEM, (Singh, 1995) the researcher's system of hypothesized equations (with multiple dependent variables) is simultaneously examined using multiple datasets – in this setting data from different countries or culture groups (Drasgow and Kanfer, 1985).

Ambiguity about construct equivalence is reduced as multiple-group SEM provides overall model and multivariate goodness-of-fit statistics and does so with flexibility. The researcher may correct for unequal reliabilities across groups (the silent threat to validity in IB research according to Davis, Douglas, and Silk, 1981), local or emic indicators can be included, and using restricted models, models can be compared across datasets with some control on overall error rate. Though it requires smaller sample sizes than IRT it is still sample hungry and may not be available for small sample studies.

Conclusion

To have similarity to compare phenomenon (the etic dimension) and significant variance to infer inter-cultural differences (the emic dimension) the cross-cultural researcher must have functional and conceptual equivalence of the examined phenomenon or construct. The rubric for cross-national, cross-cultural research now means that researchers must design construct equivalence and validity into their studies from the beginning and assess it at the finish line in order to have confidence in their conclusions about the effect of culture on a construct or relationship.

**David McArthur, Ph.D. (University of South Carolina) is an Assistant Professor of International Business and Strategy at Utah Valley State College. He also holds a BS in Marine Engineering from the US Merchant Marine Academy and an MBA and an MA in International & Area Studies (Japan) from Brigham Young University. Dr. McArthur's current research interests include understanding how firms transfer knowledge across international (and cultural) boundaries, how managers resolve problems implementing technology strategies, and how they negotiate for support and resources needed in filling their subsidiaries' roles within multinational enterprise.*

References

Adler, N.J. 1983. *Cross-cultural management research: The ostrich and the trend.*

Bagozzi, Richard P., Yi, Youjae, & Phillips, Lynn W. 1991. Assessing construct validity in

organizational research. *Administrative Science Quarterly*, 36(3): 421-458.

Berry, John W. 1969. On cross-cultural comparability. *International Journal of Psychology*, 4: 119-128.

Berry, J.W. 1980. Introduction to methodology. In H.C. Triandis & J.W. Berry (Ed.), *Handbook of cross-cultural psychology*, vol. 2: 1-28. Boston: Allyn and Bacon.

Bhagat, Rabi S., & McQuaid, Sara J. 1982. Role of subjective culture in organizations: A review and directions for future research. *Journal of Applied Psychology*, 67(5): 653-685.

Boyacigiller, Nakiye A., & Adler, Nancy J. 1991. The parochial dinosaur: Organizational science in a global context. *The Academy of Management Review*, 16(2): 262-290.

Brinberg, D., & McGrath, J.E. 1985. *Validity and the research process.* Newbury Park, CA: Sage.

Brislin, R.W. 1973. *Cross-cultural research methods.* New York: J. Wiley.

Brislin, R.W., Lonner, W.J., & Thorndike, R.M. 1980. Translation and content analysis of oral and written materials. In H.C. Triandis & J.W. Berry (Ed.), *Handbook of cross-cultural psychology*, vol. 2: 389-443. Boston: Allyn and Bacon.

Cairns, David. 1997. The future shape of harmonization: A reply. *European Accounting Review*, 6(2): 305-349.

Candell, Gregory L., & Hulin, Charles L. 1986. Cross-language and cross-cultural comparisons in scale translations: Independent sources of information about item nonequivalence. *Journal of Cross-cultural Psychology*, 17(4): 417-440.

Cohen, Jeffrey R., Pant, Laurie W., & Sharp, David J. 1996. A methodological note on cross-cultural accounting ethics research. *International Journal of Accounting*, 31(1): 55-66.

- Cohen, Jeffrey, Pant, Laurie, & Sharp, David. 1993. A validation and extension of a multidimensional ethics scale. *Journal of Business Ethics*, 12(1): 13-26.
- Cook, T., & Campbell, D.T. 1976. The design and conduct of quasi-experiments and true experiments in field settings. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*: 223-326. Chicago: Rand McNally College Pub. Co.
- Davis, H.L., Douglas, S.P., & Silk, A.J. 1981. Measurement unreliability: A hidden threat to cross-national marketing research?. *Journal of Marketing*, 45(2): 98-109.
- Dean, Graeme, & Clarke, Frank. 2005. 'True and fair' and 'fair value'—accounting and legal will-o'-the-wisps. *Abacus*, 41(2): 1-9.
- DeVera, M. 1986. *Establishing cultural relevance and measurement equivalence using emic and etic items* (Ph.D. dissertation) University of Illinois at Urbana-Champaign.
- Douglas, S.P., & Craig, C.S. 1983. *International marketing research*. Englewood Cliffs N.J.: Prentice-Hall.
- Drasgow, Fritz, 1987. Study of the measurement of the measurement bias in two standardized psychological tests. *Journal of Applied Psychology*, 72, 19-29.
- Drasgow, Fritz, & Kanfer, Ruth. 1985. Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology*, 70(4): 662-680.
- Dubinsky, Alan J., Michaels, Ronald E., Kotabe, Masaaki, Lim, Chae U., & Moon, Hee-Cheol. 1992. Influence of role stress on industrial salespeople's work outcomes in the United States, Japan, and Korea. *Journal of International Business Studies*, 23(1): 77-99.
- Durvasula, Srinivas, Netemeyer, Richard G., Andrews, J. C., & Lysonski, S. 2006. Examining the cross-national applicability of multi-item, multi-dimensional measures using generalizability theory. *Journal of International Business Studies*, 37(4): 469-483.
- Flower, John. 1997. The future shape of harmonization: The EU versus the IASC versus the SEC. *European Accounting Review*, 6(2): 281-304.
- Flower, John. 1998. The future shape of harmonization: A reply. *European Accounting Review*, 7(2): 331-334.
- Frijda, Nico, & Jahoda, Gustav. 1966. On the scope and methods of cross-cultural research. *International Journal of Psychology*, 1(2): 109-127.
- Glaum, Martin. 2000. Bridging the GAAP: The changing attitude of German managers towards Anglo-American accounting and accounting harmonization. *Journal of International Financial Management & Accounting*, 11(1): 23-25.
- Hulin, Charles L. 1987. A psychometric theory of evaluations of item and scale translations: Fidelity across languages. *Journal of Cross-cultural Psychology*, 18(2): 115-142.
- Hulin, Charles L., Komocar, John, & Drasgow, Fritz. 1982. Applications of item response theory to analysis of attitude scale translations. *Journal of Applied Psychology*, 67(6): 818-825.
- Jacoby, Jacob. 1978. Consumer research: A state of the art review. *Journal of Marketing*, 42(2): 87-96.
- Kotabe, Masaaki. 1992. A comparative study of U.S. and Japanese patent systems. *Journal of International Business Studies*, 23(1): 147-168.
- Lincoln, James R., & Kalleberg, Arne L. 1985. Work organization and workforce commitment: A study of plants and employees in the U.S. and Japan. *American Sociological Review*, 50(6): 738-760.

- Lytle, Anne L., Brett, Jeanne M., Barsness, Zoe I., Tinsley, Catherine H., & Janssens, Maddy. 1995. A paradigm for confirmatory cross-cultural research in organizational behavior. *Research in Organizational Behavior*, 17: 167-214.
- Mullen, Michael R. 1995. Diagnosing measurement equivalence in cross-national research. *Journal of International Business Studies*, 26(3): 573-596.
- Nasif, Ercan G., Al-Daeaj, Hamad, Ebrahimi, Bahman, & Thibodeaux, Mary S. 1991. Methodological problems in cross-cultural research: An updated review. *Management International Review*, 31(1): 79-91.
- Osgood, C.E., May, W.H., & Miron, M.S. 1975. *Cross cultural universals of affective meaning*. Urbana, IL: University of Illinois Press.
- Peng, T. K., Peterson, Mark F., & Shyi, Yuh-Ping. 1991. Quantitative methods in cross-national management research: Trends and equivalence issues. *Journal of Organizational Behavior*, 12(2): 87-107.
- Pike, K.L. 1967. *Language in relation to a unified theory of the structure of human behavior*, (2d, rev. ed.). The Hague: Mouton.
- Przeworski, Adam, & Teune, Henry. 1966. Equivalence in cross-national research. *The Public Opinion Quarterly*, 30(4): 551-568.
- Sears, Robert R. 1951. A theoretical framework for personality and social behavior. *American Psychologist*, 6(9): 476-482.
- Sekaran, Uma. 1983. Methodological and theoretical issues and advancements in cross-cultural research. *Journal of International Business Studies*, 14(2, Special Issue on Cross-Cultural Management): 61-73.
- Sharma, Subhash, & Weathers, Danny. 2003. Assessing generalizability of scales used in cross-national research. *International Journal of Research in Marketing*, 20(3): 287-296.
- Singh, Jagdip. 1995. Measurement issues in cross-national research. *Journal of International Business Studies*, 26(3): 597-619.
- Straus, Murray A. 1969. Phenomenal identity and conceptual equivalence of measurement in cross-national comparative research. *Journal of Marriage and the Family*, 31(2, Cross-Cultural Family Research): 233-239.
- Triandis, H.C. 1972. *The analysis of subjective culture*. New York: Wiley-Interscience.
- Triandis, Harry C., & Marin, Gerardo. 1983. Etic plus emic versus pseudoetic: A test of a basic assumption of contemporary cross-cultural psychology. *Journal of Cross-cultural Psychology*, 14(4): 489-500.
- Van Raaij, W. Fred. 1978. Cross-cultural research methodology as a case of construct validity. In K.H. Hunt (Ed.), *Advances in consumer research*, vol. 5: 693-701. Urbana IL: Association for Consumer Research.
- Venkatraman, N., & Grant, John H. 1986. Construct measurement in organizational strategy research: A critique and proposal. *The Academy of Management Review*, 11(1): 71-87.
- Wallace, W.L. 1971. *The logic of science in sociology*. Chicago, IL: Aldine Atherton.
- Werner, O., & Campbell, D.T. 1970. Translating, working through interpreters, and the problem of decentering. In R. Naroll & R. Cohen (Ed.), *Handbook of methods in cultural anthropology*: 398-420. New York: Natural History Press.