

Tilburg University

## Methodological issues in psychological research on culture

van de Vijver, F.J.R.; Leung, K.

*Published in:*  
Journal of Cross-Cultural Psychology

*Publication date:*  
2000

*Document Version*  
Peer reviewed version

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
van de Vijver, F. J. R., & Leung, K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, 31(1), 33-51.

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Methodological Issues in Psychological Research on Culture

Fons J. R. van de Vijver

Tilburg University

&

Kwok Leung

City University of Hong Kong

## Abstract

The extent to which methodological tools can help correct the overemphasis on fact-finding and speed up the slow theoretical progress in cross-cultural psychology is analyzed. Two types of contributing forces of the current predicament are delineated. First, cross-cultural psychologists have created their own partis pris (such as the uncritical acceptance of cross-cultural differences in the social domain, and the uncritical rejection of such differences as biased in the cognitive domain). Second, partis pris have been inherited from mainstream psychology, such as the paradigmatic organization of research (e.g., individualism—collectivism). In the future most cross-cultural studies will be carried out by researchers who have an interest in cultural variations on a specific variable or instrument (“sojourners”), while the group of “natives” who spend most of their professional life in cross-cultural psychology will remain small but influential. Methodological issues arising in studies by both groups are described. Important trends are (a) the change from the exploration to the explanation of cross-cultural differences, which has implications for the design of cross-cultural studies, and (b) the, so far hesitant usage of recently developed statistical techniques, such as item response theory, structural equation modeling, and multilevel modeling.

## Methodological Issues in Psychological Research on Culture

Academic psychology exists now for more than a century, but complaints about the sluggish progress are not hard to find in the literature. The slow progress is astonishing if it is realized how much empirical work has been carried out in the last century. Despite the recent development of meta-analysis as a tool to integrate the findings of independent studies (e.g., Hedges & Olkin, 1985), it has turned out to be very difficult to build up a body of replicable, validated knowledge of human behavior. Cross-cultural psychology, one of the younger sprouts of the family, shows the same slow development. That the development is slow may not be immediately appreciated; after all, interest in the field has increased, as manifested in the large and consistent growth in number of publications on cross-cultural similarities and differences that appear each year. When comparing the first and second editions of the *Handbook of Cross-Cultural Psychology*, published in 1980 and 1997 respectively, one may indeed be struck by the vibrant activity in the field in the last decades. This growth, however, is largely due to the pursuit of new interests and areas, rather than the systematic accumulation of knowledge and productive paradigmatic shifts. For instance, whereas in the first edition there was still quite some attention to cross-cultural applications of Piagetian experiments, in the second edition there is more emphasis on “contextualized cognition” (e.g., Schliemann, Caraher, & Ceci, 1997). Perhaps the largest increases are found in social psychology of the self and, related to it, individualism—collectivism (Triandis, 1994; Kagitcibasi, 1997). These changes in research focus all accurately reflect the

dynamics of the research field, but do not represent changes prompted by Popperian critical experiments that demonstrated the invalidity of old theories.

The present article addresses the question to what extent methodological tools can help to overcome the poor cumulative nature of cross-cultural research. It is not our intention to deny the value and role of theories and models in advancing cross-cultural psychology; quite the contrary, we attempt to find ways in which methodological and statistical tools can help to develop testable theories and models in cross-cultural psychology.

In the first part impediments to progress in cross-cultural psychology are described. They all derive from what could be called the partis pris (preconceived opinions, prejudices) of cross-cultural psychologists. We prefer the French term instead of “prejudice” that is the common term in psychology in order to make it clear that we mean prejudices held by cross-cultural psychologists (as opposed to prejudices studied by them).

A taxonomy of cross-cultural studies (Van de Vijver & Leung, 1997a, b) is presented in the second part as they are characterized by different methodological and statistical issues. Four types of cross-cultural studies are distinguished; they are either exploratory or hypothesis testing, and involve or do not involve contextual information about the participants.

In order to appreciate the demands to be imposed on methodology and statistics in the coming decades in cross-cultural psychology, cross-cultural researchers are divided in two groups: “natives” and “sojourners.” The former direct much or even all their research effort to cross-cultural topics. Membership directories of associations that focus on cross-cultural research

like the International Association for Cross-Cultural Psychology and its French-language counterpart “Association pour la recherche interculturelle” provide many examples. In the context of this article “sojourners” refer to persons who have their primary expertise in another content domain, and who attempt to extend their research effort to different cultural groups. The latter group is responsible for a large majority of cross-cultural publications. Their needs and impact should not be dismissed. Therefore, the third part of the paper explores the methodological and statistical issues for both types of researchers separately. Our position is that by integrating substantive, methodological, and statistical issues, the validity and replicability (reliability, generalizability) of research findings should greatly increase. What we see as important future trends are presented in the last section.

#### Impediments to Progress in Cross-Cultural Psychology

No book about cross-cultural psychology is complete without a section or chapter on prejudices and cultural biases in judgments of other cultures. These prejudices work like cognitive schemata that have a bearing on the type of processing that takes place; information congruent with the schema tends to be more actively sought after and better remembered. These schemata essentially act as templates and reduce the rich and pluriform reality to more manageable formats. Cross-cultural psychologists also have such templates. From a practical perspective, these templates function like preconceptions, leading cross-cultural researchers to focus on convergent evidence and keeping them from exploring new evidence from a more neutral vantage point.

These partis pris come from two sources. Some were inherited from the parent disciplines, notably mainstream psychology, and some were created as cross-cultural psychology developed. Examples of partis pris that have gradually emerged in the discipline include:

- Uncritical acceptance of observed differences in the social domain as reflecting valid cross-cultural differences (cf. Faucheux, 1976): Öngel and Smith (1994) have shown that social psychology is the most popular domain in publications in the Journal of Cross-Cultural Psychology. The field is obviously investing much effort in documenting cross-cultural differences in social behavior, but this research often misses a critical reflection on the nature of these differences. Some social psychological differences may be deeply rooted in a culture and may even manifest themselves in a wide variety of behaviors, whereas others merely reflect superficial conventions that are unrelated to other psychological processes.
- Uncritical rejection of observed cross-cultural differences in the cognitive domain as measurement artifacts: It is interesting to note that in mental testing a view opposite to the one held in social psychology is dominant. Differences in performance across different cultural groups are often seen as due to fallacies of the tests. The often held implicit assumption seems to be that a culturally unbiased mental test should not show cross-cultural differences. This view is difficult to reconcile with the also quite popular view that cross-cultural differences in socialization practices are substantial and have an impact on many areas of psychological functioning.

- Insufficient attention for equivalence and bias: In psychology we tend to pay little attention to sampling procedures. The replicability of our results would improve if we would follow the sampling procedures that are the standard in cross-national survey research (cf. Kish, 1965). The implementation of sampling procedures will cost time and money, but the return on investment can be huge, because unwanted sample differences are less likely to confound the differences (see Leung, Lam, & Lau, 1998, for an example of how to disentangle the effects of cultural and demographic variables). Analogously, in the last decades various statistical techniques have been proposed to scrutinize bias in measures (e.g., Van de Vijver & Leung, 1997a, b), but these techniques are only infrequently applied in cross-cultural psychology, despite their obvious relevance.
- Overgeneralizations: Our samples are often small and chosen more for convenience than appropriateness; similarly, our instruments are often short and do not adequately cover the underlying construct or behavior domain of interest (Embretson, 1983). These factors alone or in combination lead to a poor replicability of results, and this sub-optimal mapping of constructs may be one reason for seemingly conflicting results reported by different researchers. Research on how to tackle this problem is not popular in cross-cultural psychology; some exceptions include the definition of the domain of generalization (indicating which psychological construct the test can be assumed to cover well) (cf. Van de Vijver & Poortinga, 1982), and the representativeness of a sample (indicating to what populations or parts thereof the results can be generalized on the



basis of the sampling scheme used). As a consequence, we are too easily inclined to conclude that there are real cross-cultural differences (i.e., differences related to the target construct) where a simple alternative explanation (like biased sampling) exists.

Problems in the accurate interpretation of score differences in cross-cultural psychology are further aggravated by what could be called the "interpretation paradox" of cross-cultural differences: Cross-cultural score differences that are larger and easier to observe and replicate, are more difficult to interpret. The largest cross-cultural differences tend to be observed between culturally highly distinct populations. If a mental test like the Raven Matrix is administered to both literate and illiterate groups, the pattern of outcomes is predictable and replicable; but would it be easily interpretable? The real interpretation issues emerge when we attempt to go beyond the rather empty statement that the scores differ because the groups have different cultural backgrounds (for simplicity of the argument, we omit possible differences in genetic factors) and explore more precise explanations of the score differences. The cultural backgrounds of literate and illiterate subjects differ in so many respects (e.g., education, socialization, daily experiences, and exposure to media), that there is a problem of identifying the real cause. The number of rival explanations tends to increase with the cultural distance. Thus, the interpretation paradox of cross-cultural differences holds that score differences found in closely related cultures may be relatively hard to find but once reliably identified, easy to interpret; score differences as found in widely diverging cultures, are relatively easy to interpret, but they tend to be open to multiple interpretations.

From the perspective of the history of science it is interesting to observe that cross-cultural psychology has been rather successful in conveying the message that psychological constructs like personality, emotion, cognition, and social behavior cannot be studied in a cultural vacuum. It is reassuring to see that new textbooks in developmental, social, and personality psychology tend to pay more and more attention to cultural factors. From this perspective the rebellion against our parents has been successful. At the same time, however, we have inherited some of their bad habits. Progress in cross-cultural psychology has also been hampered by the adoption of partis pris of mainstream psychology:

- Paradigmatic organization of research: Kuhn (1970) has described how scientists, not just psychologists, tend to organize their research in accordance to dominant paradigms. In their initial stage paradigms facilitate research and hypothesis testing, but eventually perish because they are insufficiently self-critical and cannot generate fresh new theories and instruments that may be needed to overcome typically well-documented problems. Probably the best example in cross-cultural psychology has been cognitive style research (Berry, Poortinga, Segall, & Dasen, 1992); a more recent example is individualism—collectivism (e.g., Kagitcibasi, 1997; Triandis, 1995). The uncritical usage of such dimensions to explain cross-cultural differences in various psychological constructs has major drawbacks. First, the validity of a dimension like individualism—collectivism employed in the explanation of cross-cultural differences is often not demonstrated; it is disturbing to observe how infrequently the concept is actually measured and how often it is merely

utilized as a post hoc explanation. Second, much of the research is based on two-country comparisons, often a South-East Asian country and the USA. Apart from an assumed cultural homogeneity within these countries, there is the issue that between these societies differences exist in many respects: economical, sociological, political, to mention just a few.

Alternative explanations are often not explored. Individualism—collectivism is now used so extensively in cross-cultural psychology that it is easy to predict that it will soon lose its attractiveness for reasons that have caused other once dominant paradigms to lose their followers: an abundance of data that cannot be reconciled with the theory, a proliferation of definitions, or simply loss of interest in the topic by researchers (Meehl, 1991). That the individualism—collectivism dimension will perish under its own success would be a pity. At the country level individualism—collectivism shows a high positive correlation with Gross National Product (Hofstede, 1980). Therefore, the dimension could help us to better delineate the psychological consequences of economic development.

- Focus on significance testing, insufficient usage of effect size estimates, and scant attention to pattern differences: Most cross-cultural psychologists, like mainstreamers, have been brought up in the classical Neyman—Pearson framework, in which an experimental and a control group, which differ in one outcome-relevant aspect only, are tested for differences with regard to some outcome variable. The framework that focuses on statistical significance, is known not to function well in cross-cultural psychology where groups to be compared never differ in one

aspect only (Poortinga & Malpass, 1986). The framework has also come under fire in mainstream psychology (Cohen, 1990); the critique focuses on various aspects, such as the “reification” of the .05 level of significance, problems of getting results published that are not significant at the .05 level, and the insufficient realization that the actual significance level is a combination of size of the sample and the cultural differences. Cohen has repeatedly argued that it would be useful to add at least effect size estimates to common significance tests. Such effect sizes provide an indication of how far are groups apart (in terms of their pooled standard deviation). In addition, effect sizes are easier to compare than significance levels. If one is interested in comparing cultural groups on sets of variables (such as the items of an instrument) instead of a single one, effect sizes provide good input for an analysis in which the aim is for an examination of profiles and/or patterns of variables instead of a single variable.

- Poor measurement of the environment/social context: The quality of our instruments to measure individuals exceeds by far the quality of the measurement of the environment. This is hardly surprising in mainstream psychology, which after all deals more with individuals than with their environments, but the prominence of this one-sided development is more surprising in cross-cultural psychology. For a science that deals with human—environment interactions, it is indispensable to develop instruments to assess the environment, both physical and social.
- Western bias: Many writers have lamented about the Western bias in psychology in general, and cross-cultural psychology in particular (e.g.,

Sinha, 1987). The bias is reflected in the methods used, the theoretical orientations adopted, and the topics chosen for study. For instance, there has been severe criticism of validity and reliability problems associated with a blind importation of Western instruments in non-Western countries (e.g., Cheung, 1996). Ho (1998) complained about the wide acceptance of methodological individualism in non-Western countries, which does not capture the essence of their social reality. Moghaddam (1990) noted that research topics in non-Western countries are often dominated by research trends in the West. In response to these problems, Leung and Zhang (1995), after a review of various viewpoints, concluded that indigenous research and theorizing as well research that integrates different cultural perspectives are crucial to the establishment of more useful and universal psychological theories. A good example is the work on Chinese personality by Cheung, Leung, and their colleagues, who adopted a completely indigenous approach. Their work has suggested that the Big-Five model of personality is incomplete for Chinese because of the identification of a sixth factor, the so-called Chinese tradition factor. (Cheung et al., 1996). Current work shows that this Chinese factor is identifiable in an American sample from Hawaii, which suggests that it may not be a culture-specific factor of personality (Cheung et al., 1998).

#### A Taxonomy of Cross-Cultural Studies

In Table 1 four types of cross-cultural studies are distinguished, based on two underlying dimensions (cf. Van de Vijver & Leung, 1997a, b). The first one refers to the common distinction between exploratory and hypothesis-testing studies (e.g., Christensen, 1997); some studies set out to explore

cross-cultural differences without strong prior ideas about where to expect these, while other studies are guided by theoretical frameworks or earlier results that enable the formulation of a priori hypotheses. The second dimension is more specific to cross-cultural research; a distinction is made between studies with or without the consideration of contextual factors. We briefly discuss the four cells of the table.

In generalizability studies there is (a) a strong theoretical framework, that allows for the formulation of hypotheses about cross-cultural differences and similarities, and there is (b) no measurement of contextual factors. Schwartz's (1992) work on values, Eysenck's work on the universality of his three-factorial personality structure (Eysenck & Eysenck, 1983), and McCrae and Costa's (1997) work on the universality of the five-factor model of personality are good examples. In all these studies there is an emphasis on the universality of particular structures and there is little concern for identifying or measuring contextual factors as potentially confounding factors.

Theory-driven studies share this strong theoretical background; however, here contextual information is also utilized. Berry's (1976) work on cognitive style provides a good example. Berry assumed that visual discrimination, visual disembedding, and spatial orientation should be more important for survival for hunters and gatherers than for members of agricultural groups. Therefore, hunters and gatherers should be more field-independent and members of agricultural groups more field-dependent (field independence is the tendency to use internal frames of reference such as bodily cues to orient oneself in space, whereas field dependence is the tendency to use external frames). Berry's study confirmed that members of

hunting groups were indeed more field independent than were members of agricultural groups.

The third type of study is not theory-based and does not consider contextual factors. It is by far the most common type of research; many studies reported in the Journal of Cross-Cultural Psychology and the International Journal of Psychology in the 70s and 80s fall into this category. In many cases an instrument that has been found to show adequate properties in one cultural group is translated and administered in a new cultural context. The study is undertaken to get some insight in the cross-cultural stability of the structure found in Western groups (e.g., is neuroticism the same concept around the globe?) or to compare the scores obtained in cultural groups (e.g., is there a difference in average level of neuroticism between countries A and B?).

The fourth and last type is called external validation. A recent example can be found in the work by Williams, Satterwhite, and Saiz (1998). They asked persons in 20 countries to indicate the psychological importance of 300 psychological traits from the Adjective Check List. Their analysis mainly focuses on country comparisons of average scores. These averages were correlated with various country characteristics, such as affluence and population density. It was found that affluence showed a strong relationship with psychological importance (more affluent countries tend to show lower scores), while population density did not show any relationship. Georgas, Van de Vijver, and Berry (1999) examined the relationship of religion, affluence, and various psychological indicators of countries. Their most important overall conclusion was religion and affluence appeared to have opposite

psychological effects. For most religions, higher proportions of followers were associated with more emphasis on interpersonal aspects, such as power, loyalty, and hierarchy, whereas more affluence tends to lead to more emphasis on intra-personal aspects, such as individualism, utilitarian commitment, and well-being.

In order to compare the designs of the studies, it is important to briefly discuss their strengths and weaknesses (see Table 2). Generalizability studies tend to pay ample attention to equivalence and bias issues, thereby guarding themselves against claims of poor measurement. Their most important weakness is the absence of any contextual variable that could shed light on the nature of cross-cultural differences observed. As long as such studies are interested in the universality of a particular structure (e.g., of intelligence or values), the absence of contextual information is not a problem.

Theory-driven studies examine the relationship of cultural factors and behavior, which is seen by many as the core of cross-cultural psychology. Their explicit focus on questions that are so central to cross-cultural psychology is their main asset. In practice, their weakness may be lack of attention to alternative interpretations. These studies tend to focus on a single explanation, thereby possibly neglecting alternative interpretations.

The major advantage of psychological differences studies is their “open-mindedness” about cross-cultural differences. Their broad scope on cross-cultural differences makes them suitable for exploring cross-cultural in under-searched domains. Their virtue can easily become their Achilles heel, however, because their openness usually does not help the researcher in the



interpretation of the differences (e.g., would the same differences be found when using other instruments?).

It is an asset of external validation studies that they focus on exactly the weakness of psychological differences studies: the interpretation of cross-cultural differences is the aim of such studies. A likely problem in these studies is the choice of country characteristics to which the psychological variables can be related. Hundreds of country characteristics are available nowadays; the Internet, in particular pages of large international bodies such as the United Nations, the World Health Organization, and the World Bank, provide rich sources for country-level data. With so many variables available and so little theory to pinpoint the relevant ones, there is an inevitable problem of choice and selection. The selection issue is compounded by the often strong intercorrelations of the indicators. Georgas et al. (1999) studied indicators from various domains of more than 100 countries: ecology (e.g., precipitation rate), economy (e.g., Gross National Product), education (e.g., enrollment ratios at primary, secondary, and tertiary level), mass communication (e.g., number of newspapers), and population (e.g., infant mortality). The more than 20 indicators that they examined showed a strong first factor in a factor analysis, labeled affluence by the authors. At first sight this finding seems attractive because of its parsimony; there is apparently no need to measure many variables to obtain an adequate, stable country score on this single factor. Yet, paradoxical as it may sound, the choice of these few indicators can be problematic. Suppose two researchers are interested in the explanation of cross-cultural differences in cognitive test performance. Researcher A opts for the measurement of ecological indicators at country

level to explain cross-national score differences, while researcher B prefers educational indicators. Both researchers may well find that their country variables are effective predictors of country score differences. Who is right? Because of the high intercorrelations of ecological and educational indicators, it is likely that in statistical analyses both sets of predictors are interchangeable. It is a prudent strategy to refer to the general underlying factor, affluence, instead of to one of the specific clusters that make up the general factor.

#### Development in Research in the Near Future

We contend that methodological developments will be somewhat different for “sojourners” and “natives.” To some extent both types of research will show their own dynamics. This difference may be primarily due to differences in the types of research questions studied in the two traditions. “Sojourners” will be mainly interested in two types: (a) psychological differences studies - their interest in cross-cultural differences will be mainly exploratory; and (b) generalizability studies – they are eager to show the universality of their theoretical propositions. However, it should be noted that they tend not to work on theories that capture the patterning of cross-cultural similarities and differences. Their aim is more modest, namely an exploration of cross-cultural similarities and differences in a specific domain. In contrast, “natives” will have more interest in culture per se, and it can be expected that they become increasingly interested in questions covering broad, domain-transcending areas of the field, such as the formulation of theories of cross-cultural similarities and differences.

### Research Developments for “Sojourners”

Forces that gave momentum to the current interest in cross-cultural studies, such as the globalization of the market and an increase in cross-cultural encounters in daily life, will continue to be significant (Marsella, 1998). It can be expected that the current strong interest in cross-cultural studies will continue to contribute to further growth of the field.

With the advent of cross-cultural studies it can be expected that the standards of these studies will increase. It will become increasingly difficult for “sojourners” to publish the “safari” type of research, in which a test is administered in two highly dissimilar groups, and the averages are compared (without any concern for the suitability of the instrument and the equivalence of the scores). Editors and reviewers will become increasingly aware of the specific demands of cross-cultural research.

Cross-cultural psychologists have an important task in communicating these standards to the general field. It is important that we communicate how important it is to deal with alternative interpretations in cross-cultural psychology, how we can study equivalence, how we can compare nonequivalent groups, etc. The American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education have jointly formulated standards for developing and administering psychological and educational tests. Even though they have been formulated with the USA as their primary area of application, they can be transferred to various other places without many adaptations. It is important that we develop and disseminate a similar set of standards. A first attempt can be seen in the initiative of the International Test Commission, in

which Ron Hambleton headed an international committee to formulate guidelines (recommended practices) about appropriate test translations and adaptation (Hambleton, 1994; Hambleton & Van de Vijver, 1996).

Sojourners who are interested in testing their theories in different cultural milieus will also increase. It is important that their effort is connected to development in cross-cultural psychology, and a successful integration will definitely benefit “sojourners” and “natives,” and bring about significant progress in the field. For instance, in an impressive program of research, Cohen and Nisbett (Cohen & Nisbett, 1993; Nisbett, 1993) have shown that ecology is related to the development of a code of honor, which affects how personal affront is dealt with. This work connects well with earlier work on ecology and cognitive style, although these two programs of work cover different domains (cognitive vs. social).

#### Research Development for “Natives”

A first development refers to the interactions of individual and cultural factors. Traditionally the impact of cultural factors on the individual has been emphasized in cross-cultural psychology. Although the influence of individual actions on culture is beyond dispute, this line of influence has not been studied extensively. In other words, individual and cultural factors have not been frequently studied in terms of their interaction. In cultural psychology the interaction gets a more central place; it is argued there that individual and culture make each other up (e.g., Cole, 1997; Miller, 1997). Unfortunately, the approach does not yet specify clear methodological guidelines as to how the interaction can be studied. It is expected that in the coming decades more

advanced interaction models will be developed with a bearing on how (cross-)cultural research should be carried out.

The second type of developments for “natives” stems from methodological and statistical innovations. It is astonishing to observe how frequently new statistical tools that can address previously intractable questions are ignored in current cross-cultural research. An example is Item Response Theory (e.g., Hambleton & Swaminathan, 1985; Hambleton, Swaminathan, & Rogers, 1991). In translation projects it may turn out that some items can be literally translated while others need to be adapted (i.e., the item content has to be changed) in order to be suitable for the new cultural context. Common statistical techniques, such as exploratory factor analysis,  $t$  test, and analysis of variance, do not allow for the joint examination of the common (i.e., unchanged) and adapted (i.e., different) items. Now, assuming that all items reflect the same underlying construct, Item Response Theory allows for the comparison of scales even when not all items are identical in the groups compared. Analogously, the estimated item parameters (comparable to item means in more conventional analyses) do not depend on the specific sample of respondents. Whereas item means depend on sample particulars in conventional analyses (the same item can have a low mean in one group and a high mean in another), the estimated item parameters in Item Response Theory are independent of the score level of a group. This characteristic is useful in examining the equivalence of translations; when working with groups of monolinguals, there is no need to match them on the underlying construct in order to compare item parameters across languages (Ellis, Becker, & Kimmel, 1993). The major limitation of

Item Response Theory lies in the fairly large sample sizes that are needed. Rules of thumb often refer to at least 250 individuals per cultural group.

A second technique that is insufficiently exploited in cross-cultural research is Structural Equation Modeling (Byrne, 1989, 1994), a summary label for various statistical techniques, such as confirmatory factor analysis and path analysis with or without latent variables. Structural equation Modeling has at least two attractive features. First and foremost, it enables a fine-grained analysis of equivalence. Confirmatory factor analysis allows for detailed and highly informative comparisons of factor models across cultural groups (e.g., Cheung & Rensvold, in press; Cudeck & Claassen, 1983). All relevant parameters of such models, such as factor loadings, factor correlations, and error variances of items, can be tested for equality across groups. Second, structural equation modeling allows for a comparison of latent means. Instead of comparisons of observed scores, which may be influenced by bias, structural equation modeling allows for a comparison of subscales for which equivalence has been shown (e.g., Little, Oettingen, Stetsenko, & Baltes, 1995). The most important limitation of structural equation modeling is the problem of model fit (Bollen & Long, 1993). Dozens of fit measures have been developed and only slowly there is some consensus growing about which fit statistics are useful for which purposes.

The third type of technique with a large potential value for cross-cultural research is multilevel modeling (Bryk & Raudenbush, 1992). In cross-cultural psychology we often come across research in which we want to compare findings at the individual and cultural level (Leung, 1989; Leung & Bond, 1989). It is claimed in multilevel research that results may differ across

levels of aggregation. For example, subjective well-being is positively correlated with national income at the national level (more affluent countries report on average higher levels of well-being), while no such relationship exists when data were examined for the US during the last decades (Myers & Diener, 1996). Subjective well-being did not increase in this period, despite the sizable net increase in income during this period. As another example, Entwistle, Mason, and Hermalin (1985) studied the relationship between socioeconomic status and fertility, and found a relationship between a country's affluence and the average number of children born. However, in less affluent countries there tends to be a positive relationship between status and fertility, whereas in more affluent countries a negative relationship is often found.

Multilevel research addresses two types of questions that are both relevant in cross-cultural research. As a (hypothetical) example of the first type, suppose that a researcher wants to examine (intra-national) individual and cross-national differences in individualism—collectivism. He or she administers a questionnaire to persons in various countries. In addition to these data, background information on income and level of schooling is gathered. At the country level, indicators of affluence (say, Gross National Product and income inequality) are measured. In a multilevel analysis, the first step involves the analysis of the individual-level data - the background data are regressed on the individualism—collectivism scores. The second step also includes the country-level data and the regression coefficients of the first analysis are explained on the basis of the affluence indicators. This type

of analysis addresses differences in score levels across individuals within a single country as well as across countries.

The second type of analysis is aimed at a comparison of structures across aggregation levels. It involves the question as to whether a particular construct has the same meaning at individual and national level. Triandis (1995) has argued that the individualism—collectivism dimension has a somewhat different meaning at individual and country level. He proposed to use individualism—collectivism to denote the country-level dimension and “idiocentrism—allocentrism” to refer to individual level. Muthén (1991, 1994) has developed a statistical technique, multilevel factor analysis, to compare structures across aggregation levels. Using confirmatory analysis, he compares structures across aggregation levels. Van de Vijver and Poortinga (1999) use exploratory factor analysis, followed by target rotations and the computation of an agreement index. They re-analyzed data of the 1990-1991 World Values Survey (Inglehart, 1993, 1997). The study involved a total of 47,871 respondents from 39 regions. Attitudes toward postmaterialism were measured. Postmaterialists tend to emphasize self-expression and quality of life, whereas materialists emphasize economic and physical security above all (Inglehart, 1997, p. 4). The inventory (of 12 items) is assumed to show a single underlying dimension. In the first analysis the factor analytic results of each country were compared to the results of a factor analysis on a single data set in which all samples were pooled. It was found that more affluent countries tended to show a better agreement with the pooled factor solution than less affluent countries. A comparison of the within-country data and the between-country analyses revealed that most items have a similar meaning at



the individual and country level. However, a few items seemed to show a somewhat different meaning at both levels. For example, an item about fighting crime was related to a materialist attitude in the pooled within-country data, while in the between-country data factor analysis the item was associated with a postmaterialist attitude. An inspection of the data showed that the effect was probably due to the very high scores of the former Eastern Bloc countries on the item. Compared to individuals from elsewhere, persons living in these countries expressed much concern about increasing crime rates. In these countries crime fighting is more associated with a postmaterialist attitude and quality of life than with a materialist attitude. It was concluded that within the more affluent countries the scale measures the same construct, both across individuals and countries. There is less evidence for this equivalence when less affluent countries are included, particularly because some items show a different meaning at the individual and country level.

The three relevant developments mentioned all involve advances in methodology and psychometrics. The last one does not hinge on developments in statistics and methodology; it involves the quantification of bias and equivalence. In cross-cultural research we tend to treat bias and equivalence as dichotomous phenomena: data are biased or unbiased and our statistical analyses are geared at establishing which of the two possibilities applies. This dichotomy in our understanding of bias and equivalence has impeded the advancement of our field, and there are already a number of ways to transcend this simple dichotomy and to achieve a more balanced treatment of bias and equivalence.

One approach targets at the study design by including measurement of presumably biasing factors in addition to measures of target constructs (Poortinga & Van de Vijver, 1987). For example, if differential social desirability is likely to explain cross-cultural differences in some personality or attitude questionnaire, one could include a measure of social desirability in the design to verify this possibility. As another example, in a cross-cultural study of mental test performance, one could include measures of parental characteristics or school quality in order to examine to what extent these measures may account for the cultural differences observed. An empirical example is due to Poortinga (cf. Poortinga & Van de Vijver, 1987), who studied the habituation of the Orienting Reflex among illiterate Indian tribes and Dutch military conscripts. The Skin Conductance Response, the dependent variable, was significantly larger in the Indian group. It could be argued that intergroup differences in arousal could account for these differences. Arousal was operationalized as the spontaneous fluctuations in Skin Conductance Response in a control condition. After statistically controlling for these fluctuations using a hierarchical regression analysis, the cross-cultural differences in habituation of the Orienting Reflex disappeared. The measurement of contextual factors to verify or falsify particular interpretations of cross-cultural differences can be expected to gain popularity and importance. This strategy is obviously superior to the reliance on various post hoc interpretations, which by definition are not validated. If the field of cross-cultural research gradually develops from the exploration to the explanation of cross-cultural differences, the inclusion of explanatory

variables in the design is a natural addition to the designs of psychological differences studies.

Another way of subjecting bias to empirical scrutiny is triangulation, which is implemented by a monotrait—multimethod design (Campbell & Fiske, 1959). For example, when similar cross-cultural differences are found for self-reports and peer ratings, bias is less likely to have influenced the results than when the results are obtained with a single method. When results do not converge across methods, estimates of method effects can be made, thereby providing evidence about how bias has affected the results. Thus, Hess, Chang, and McDevitt (1987) found that in comparison with American mothers, Chinese mothers were more likely to attribute the academic performance of their children to effort. Consistent with this result, Chinese children were also more likely to attribute their academic performance to their own effort than were American children. The convergence between the results of children and mothers strengthened the validity of the cultural difference observed.

Despite our firm belief that the utilization of the recently developed statistical techniques described here will augment the quality of cross-cultural research, there is no reason to expect that developments in methodology and statistics will create a revolution in measurement and theory in cross-cultural psychology. Central problems in the field, such as the conceptualization of the interaction of individual and cultural factors, are unlikely to be solved by statistical innovations. Theoretical innovations are paramount to true advances in tackling these elusive problems, while methodological and

statistical innovations may make theoretical problems more tractable and help to decide between competing theoretical positions.

### Conclusion

In the future two groups of researchers in cross-cultural psychology will gradually emerge: "natives" (emphasis on culture and the methodology for the study of culture) and "sojourners" (brief, sporadic excursions in cross-cultural research). Methodological developments will not completely coincide for the two groups. "Sojourners" will be mainly interested in psychological differences studies and generalization studies; there is a need to develop recommendations about good cross-cultural research practices and integration of monocultural and cross-cultural theorizing and findings. "Natives" will carry out research that is central to our understanding of cultural differences and the influence of culture. An important requirement in this line of research will be better usage of methodological and statistical techniques, such as Item Response Theory, Structural Equation Modeling, and multimethod designs.

The replicability of cross-cultural findings, often the Achilles heel of our empirical endeavors, will improve when we develop more sensitivity to our partis pris as cross-cultural psychologists and psychologists, when we put more emphasis on theory testing and development, and when our research is guided by appropriate methodological tools.

## References

Berry, J. W. (1976). Human ecology and cognitive style. Comparative studies in cultural and psychological adaptation. Beverly Hills, CA: Sage.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R., (1992). Cross-cultural psychology. Research and applications. Cambridge: Cambridge University Press.

Bollen, K. A., & Long, J. S. (Eds.) (1993). Testing structural equation models. Newbury Park, CA: Sage.

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis. Newbury Park, CA: Sage.

Byrne, B. M. (1989). A primer of LISREL: Basic applications and programming for confirmatory factor analytic models. New York: Springer.

Byrne, B. M. (1994). Structural equation modelling with EQS and EQS/Windows: Basic concepts, applications, and programming. Thousand Oaks, CA: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, *56*, 81-105.

Cheung, F. M. (1996). The assessment of psychopathology in Chinese societies. In M. H. Bond (Ed.), The handbook of Chinese psychology (pp. 393-411). Hong Kong: Oxford University Press.

Cheung, F. M., Leung, K., Fan, R. M., Song, W. Z., Zhang, J. X., & Zhang, J. P. (1996). Development of the Chinese Personality Assessment Inventory (CPAI). Journal of Cross-Cultural Psychology, *27*, 181-199.

Cheung, F. M., Leung, K., Zhang, J. X., Sun, H. F., Gan, Y. Q., Song, W. Z., & Xie, D. (1998). Indigenous Chinese personality constructs: Is the five-factor model complete? Manuscript under review.

Cheung, G., & Rensvold (in press). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modelling. Journal of Cross-Cultural Psychology.

Christensen, L. B. (1997). Experimental methodology (7th ed.). Boston: Allyn & Bacon.

Cohen, D., & Nisbett, R. E. (1993). Self-protection and the culture of honor: Explaining southern homicide. Personality and Social Psychology Bulletin, *20*, 551-567.

Cohen, J. (1990). Things I have learned (so far). American Psychologist, *45*, 1304-1312.

Cudeck, R., & Claassen, N. (1983). Structural equivalence of an intelligence test for two language groups. South-African Journal of Psychology, *13*, 1-5.

Ellis, B. B., Becker, P., & Kimmel, H. D. (1993). An item response theory evaluation of an English version of the Trier Personality Inventory (TPI). Journal of Cross-Cultural Psychology, *24*, 133-148.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, *93*, 179-197.

Entwistle, B., Mason, W. M., & Hermalin, A. I. (1985). The multilevel dependence of contraceptive use on socioeconomic development and family planning program strength. Demography, *23*, 199-215.

Eysenck, H. J., & Eysenck, S. B. G. (1983). Recent advances in the cross-cultural study of personality. In J. N. Butcher & C. D. Spielberger (Eds.), Advances in personality assessment (Vol. 2, pp. 41-69). Hillsdale, NJ: Erlbaum.

Faucheux, C. (1976). Cross-cultural research in experimental social psychology. European Journal of Social Psychology, 6, 269-322

Georgas, J., Van de Vijver, F. J. R., & Berry, J. W. (1999). Ecosocial indicators and psychological variables in cross-cultural research. (Manuscript under review)

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. European Journal of Psychological Assessment (Bulletin of the International Test Commission), 10, 229-244.

Hambleton, R. K., & Swaminathan H. (1985). Item response theory: Principles and applications. Dordrecht, the Netherlands: Kluwer.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Hess, R. D., Chang, C. M., & McDevitt, T. M. (1987). Cultural variations in family beliefs about children's performance in mathematics: Comparisons among People's Republic of China, Chinese-American, and Caucasian-American families. Journal of Educational Psychology, 79, 179-188.

Ho, D. Y. F. (1998). Interpersonal relationships and relationship dominance: An analysis based on methodological relationalism. Asian Journal of Social Psychology, 1, 1-16.

Hofstede, G. (1980). Culture's consequences: International differences in work-related values. Beverly Hills, CA: Sage Publications.

Inglehart, R. (1993). World Values Survey 1990-1991. WVS Program. J.D. Systems, S.L. ASEP S.A.

Inglehart, R. (1997). Modernization and postmodernization. Cultural, economic and political change in 43 societies. Princeton, NJ: Princeton University Press.

Kagitcibasi, C. (1997). Individualism and collectivism. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), Handbook of cross-cultural psychology (2nd ed., Vol. 3, pp. 1-49). Boston: Allyn & Bacon.

Kish, L. (1965). Survey sampling. New York: Wiley.

Kuhn, T. S. (1970). The structure of scientific revolutions. Chicago: University of Chicago Press.

Leung, K. (1989). Cross-cultural differences: Individual-level vs. culture-level analysis. International Journal of Psychology, 24, 703-719.

Leung, K., & Bond, M. H. (1989). On the empirical identification of dimensions for cross-cultural comparison. Journal of Cross-Cultural Psychology, 20, 133-151.

Leung, K., Lau, S., & Lam, W. L. (1998). Parenting styles and academic achievement: A cross-cultural study. Merrill-Palmer Quarterly, 44, 157-172.

Leung, K., & Zhang, J. X. (1996). Systemic considerations: Factors facilitating and impeding the development of psychology in developing countries. International Journal of Psychology, 30, 693-706.

Little, T. D., Oettingen, G., Stetsenko, A., & Baltes, P. B. (1995). Children's action-control beliefs and school performance: How do American



children compare with German and Russian children? Journal of Personality and Social Psychology, 69, 686-700.

Marsella, A. J. (1998). Toward a "global-community psychology." American Psychologist, 53, 1282-1291.

McCrae, R. R., Costa, P. T. (1997). Personality trait structure as a human universal. American Psychologist, 52, 509-516.

Meehl, P. E. (1991). Why summaries of research on psychological theories are often uninterpretable. In R. E. Snow & D. E. Wiley (Eds.), Improving inquiry in social science. A volume in honor of Lee J. Cronbach (pp. 13-59). Hillsdale, NJ: Erlbaum.

Moghaddam, F. M. (1990). Modulative and generative orientations in psychology: Implications for psychology in the three worlds. Journal of Social Issues, 46, 21-41.

Muthén, B. O. (1991). Multilevel factor analysis of class and student achievement components. Journal of Educational Measurement, 28, 338-354.

Muthén, B. O. (1994). Multilevel covariance structure analysis. Sociological Methods & Research, 22, 376-398.

Myers, D. G., & Diener, E. (1996). The pursuit of happiness. Scientific American, 274(5), 54-56.

Nisbett, R. E. (1993). Violence and U.S. regional culture. American Psychologist, 48, 441-449.

Öngel, Ü., & Smith, P. B. (1994). Who are we and where are we going? JCCP approaches its 100th issue. Journal of Cross-Cultural Psychology, 25, 25-53.

Poortinga, Y. H., & Malpass, R. S. (1986). Making inferences from cross-cultural data. In W. J. Lonner & J. W. Berry (Eds.), Field methods in cross-cultural psychology (pp. 17-46). Beverly Hills, CA: Sage.

Poortinga, Y. H., & Van de Vijver, F. J. R. (1987). Explaining cross-cultural differences: Bias analysis and beyond. Journal of Cross-Cultural Psychology, 18, 259-282.

Schliemann, A., Caraher, D., & Ceci, S. J. (1997). Everyday cognition. In J. W. Berry, P. R. Dasen, & T. S. Saraswathi (Eds.), Handbook of cross-cultural psychology (2nd ed., Vol. 2, pp. 177-216). Boston: Allyn & Bacon.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), Advances in experimental social psychology (Vol. 25, pp. 1-65). Orlando, FL: Academic Press.

Sinha, D. (1987). Psychology in India: A historical perspective. In G. H. Blowers & A. M. Turtle (Eds.), Psychology moving East: Status of Western psychology in Asia and Oceania (pp. 39-52). Boulder, CO: Westview Press.

Triandis, H. C. (1995). Individualism and collectivism. San Francisco: Westview Press.

Van de Vijver, F. J. R. & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. European Psychologist, 1, 89-99.

Van de Vijver, F. J. R., & Leung, K. (1997a). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), Handbook of cross-cultural psychology (2nd ed., Vol. 1, pp. 257-300). Boston: Allyn & Bacon.

Van de Vijver, F. J. R., & Leung, K. (1997b). Methods and data analysis for cross-cultural research. Newbury Park, CA: Sage.

Van de Vijver, F. J. R., & Poortinga, Y. H. (1982). Cross-cultural generalization and universality. Journal of Cross-Cultural Psychology, 13, 387-408.

Van de Vijver, F. J. R., & Poortinga, Y. H. (1999, in preparation). Structural equivalence in multilevel research. Tilburg University.

Williams, J. E., Satterwhite, R. C., & Saiz, J. L. (1998). The importance of psychological traits. New York: Plenum Press.

Table 1.

Types of Cross-Cultural Studies

		Orientation more on	
Consideration of contextual factors	Hypothesis testing	Exploration	
No	Generalizability studies	Psychological differences studies	
Yes	Theory-driven studies	External validation studies	

Table 2.

Major Strengths and Weaknesses of the Four Types of Studies

Type of study	Major strength	Major weakness
Generalizability studies	study of equivalence	no contextual variables included
Theory-driven studies	study of relationship of cultural factors and behavior	lack of attention to alternative interpretations
Psychological differences studies	“open-mindedness” about cross-cultural differences	Ambiguous interpretation
External validation studies	focus on interpretation	choice of covariates may be meaningless