Tilburg University

Personality in Cultural Context

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

Published in:
Journal of Personality

Publication date:
2001

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. (2001). Personality in Cultural Context: Methodological Issues. *Journal of Personality*, 69(6), 1007-1031.

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.

Personality in Cultural Context: Methodological Issues

Fons van de Vijver

Tilburg University

Kwok Leung

City University of Hong Kong

Abstract

Methodological issues in cultural and cross-cultural personality research are described. A taxonomy of these studies is presented, based on whether a study is exploratory or tests hypotheses, and whether or not contextual information is measured. Core methodological issues are bias and equivalence; a taxonomy and a brief overview of statistical procedures to examine equivalence are presented, with a focus on procedures to assessing structural equivalence (i.e., similarity of meaning of an instrument across cultures). Examples are given of studies in which cultural and cross-cultural approaches, often seen as antithetical, have been fruitfully integrated. Finally, multilevel models are described in which personality characteristics are examined at individual and cultural level.

Personality in Cultural Context: Methodological Issues

The relationship between personality and culture has been studied from various perspectives. It has become popular to refer to these by dichotomies, such as emic—etic, inside—outside view, qualitative—quantitative, culture-specific—culture-general, and cultural—cross-cultural (Berry, Poortinga, Segall, & Dasen, 1992; Poortinga, 1997). The latter dichotomy is adopted here. Studies that employ a cultural-psychological approach examine personality in a specific cultural context, whereas cross-cultural studies examine and compare personality across cultures.

The two traditions are often treated as antithetical. Studies in cultural psychology attempt to delineate constructs that optimally describe the personality of members of a particular group; unstructured or semi-structured means of data collection such as interviews are frequently employed. Cross-cultural studies aim at establishing similarities and differences in personality of cultural groups, relying on more structured means of data collection such as standardized inventories. Jahoda (e.g., 1995) has repeatedly argued that upon closer scrutiny there is no contradiction between emic and etic strategies and that both strategies are needed for the advancement of knowledge (cf. Boesch, 1996; Morris, Leung, Ames, & Lickel, 1999). We concur with his view. Without cross-cultural comparisons, psychological theory is confined to its own cultural boundaries; but a blind “exportation” of western instruments to other cultures without any concern for the appropriateness of the measures is also unlikely to lead to major theoretical advancements. A famous quote by Kluckhohn and Murray (1948; see Berry et al., 1990, pp. 186-187) states that a person is in some ways similar to all other

people, in some ways to some other people, and in some ways to no other people. The same may hold true for cultures. Some aspects of a culture are universal (such as the regulation of social sharing and power in a group; Fiske, 1991; see also Lonner, 1980), other aspects are common to conglomerates of cultures (such as hunting societies), and some aspects are unique to a culture (e.g., cultural elements regarding unique ecological features, such as the Aboriginals' and Bushmen's skills to survive in an ecologically hostile environment). Clearly, neither the cross-cultural nor the cultural approach is optimally suited to examine phenomena at all three levels.

The cultural approach is geared towards the identification of culture-specific aspects of personality. Compared to the cross-cultural approach, it imposes fewer restrictions on instruments and data (such as an a priori classification of personality factors) and provides more insight into culture-specific elements. The cross-cultural approach is deliberately more culture-general by assuming a "supercultural" vantage point (by attempting to accommodate theories and findings pertaining to two or more cultures). However, it should be noted that both approaches can address etic and emic issues. Cultural specificity is strongly supported when a cross-cultural study fails to find universal aspects (e.g., of a trait structure) and cross-validation studies have shown that this failure is not due to the particulars of an instrument or sample; analogously, universality is strongly supported by a cultural approach in which cross-cultural similarity is found, despite the differences in instruments used in the various cultures.

The dichotomy of cultural and cross-cultural psychology is often incorrectly seen as having methodological ramifications. The cross-cultural approach is then treated as quantitative and the cultural approach as qualitative. Yet, as Hsu's (1970) insightful analysis of Chinese and Americans shows, comparative studies can also be qualitative. Analogously, studies of indigenous personality can be quantitative (e.g., Cheung et al., 1996).

The remainder of the article primarily deals with quantitative approaches, which due to their compatibility with the positivist approach in psychology have received ample attention; space limitations preclude a description of qualitative methods for (cross-)cultural psychology (cf. Greenfield, 1997; Naroll & Cohen, 1970).

The present article assumes a methodological perspective on (cross-)cultural studies. The second section describes how this perspective allows us to reduce the multitude of studies in (cross-)cultural psychology to four types (cf. Van de Vijver & Leung, 1997a, b). The third section describes procedures to sample cultures. A taxonomy of bias and equivalence is also presented in the fourth section. A brief overview of statistical procedures to deal with bias and equivalence is presented in the fifth section. The sixth section describes studies in which cultural and cross-cultural approaches have been fruitfully integrated. The seventh section deals with a recent topic in personality research in (cross-)cultural psychology that can be seen as a quantitative extension of work in psychological anthropology (Bock, 1980; Piker, 1998): personality as an individual and as a cultural characteristic. Directions for future research are described in the last section.

Design of Studies in Personality Research in (Cross-)Cultural Psychology

A Taxonomy of Studies in (Cross-)Cultural Psychology

When cultural similarities and differences are evaluated, two major orientations are possible: Structure-oriented vs. level-oriented (Van de Vijver & Leung, 1997). In structure-oriented studies, the focus is on the relationship among variables, which may be correlational, ranging from simple correlations to factor structures. Good examples include whether locus of control has the same correlates in different cultures (Hui & Triandis, 1983) and whether the Big-Five Model of personality is generalizable to all cultures (McCrae, Costa, del Pilar, Rolland, & Parker, 1998). Causal relationships are possible as well, which include simple causal relationships between independent and dependent variables in experiments and complex relationships analyzed by regression analysis and structural equations modeling. For instance, Dunbar, Saiz, Stela, and Saez (2000) explored whether prejudice/tolerance predicts outgroup bias in different cultures. A more complex example involved testing whether self-esteem and relationship harmony are related to life satisfaction in the same way in Hong Kong and in the USA, and whether these two variables mediated the effects of personality traits on life satisfaction across the two cultures (Kwan, Bond, & Singelis, 1997). In contrast, level-oriented studies intend to explore whether cultures are different with regard to a certain personality dimension. For instance, Rubinstein (1996) reported that Jewish-Israeli respondents reported lower scores in Right-Wing Authoritarianism and religiosity than did Moslem-Palestinian respondents.

Van de Vijver and Leung (1997a, b) have proposed a 2×2 classification of studies in (cross-)cultural psychology based on two dimensions: Whether the purpose of the study is hypothesis-testing or

exploratory, and whether or not contextual variables are involved (see Table 1). Contextual variables are broadly defined as variables used to validate a particular interpretation of cross-cultural differences, either person-related (such as age, gender, or psychological characteristics) or culture-related (such as Gross National Product, educational systems, and health care institutions). Note that the four types of studies described are prototypes, and there are studies that may be classified into more than one type. Both structure- and level-oriented studies are possible in each of these four types of research. Both quantitative and qualitative research is also possible in each type, although quantitative work predominates in personality research.

In psychological- differences studies, the purpose is exploratory and no contextual variables are included. A typical example is the comparison of two cultures on a set of personality dimensions, which was popular in the 1970s and 1980s. This type of work is useful for its exploratory, open-minded nature, but is uninformative when it comes to the interpretation of the observed cultural dissimilarities. Donald Campbell has repeatedly argued that comparisons of two groups tend to be open to multiple interpretations (e.g., Campbell, 1986; Cook & Campbell, 1979). In generalizability studies, the focus is on the evaluation of a hypothesized cultural difference or similarity in the absence of contextual variables. A good example is the cross-cultural work on the generality of the Big-Five Model of personality in different cultures (McCrae et al., 1998). Such studies are vigorous in assessing equivalence across cultures, but may be restrictive in scope and fail to take into account emic (i.e., culture-specific) considerations.

In ecological linkage studies (termed external validation studies by Van de Vijver & Leung, 1997a, b), the purpose is to search for explanatory variables for observed cultural similarities and differences. Studies of this type are at the culture level, exploratory in nature, and include a large number of contextual variables. A good example is provided by Van Hemert, Van de Vijver, Poortinga, and Georgas (2000a), who have reported that across 30 countries, affluence, a contextual variable in this study, is negatively correlated with depression as measured by the Beck Depression Inventory. Such studies are capable of locating interesting and unexpected relationships, but like any exploratory studies, interpretation of results may be ambiguous. Finally, in contextual-theory studies (termed theory-driven studies by Van de Vijver & Leung, 1997a, b), the purpose is hypothesis-testing, and contextual variables are included as an explanation of the cultural differences observed. A good example is provided by McCrae, Yik, Trapnell, Bond, and Paulhus (1998), who found that compared to American norms, Hong Kong college students were lower in extraversion, which may be a result of the emphasis on restraint and self-effacement in Chinese culture. To evaluate this possibility, they included three groups of Chinese in Canada in a follow-up study: recent immigrants, immigrants who had been in Canada for a longer period of time, and Canadian-born Chinese. Acculturation, a contextual variable in this study, showed the expected effect, with Canadian-born Chinese showing the highest level of extraversion, followed by immigrants who had been in Canada for a long time, and then by recent immigrants. This type of studies is very useful in theory building, but may sometimes fail to evaluate competing hypotheses.

Cultural Origin of Personality Instruments

An important issue in the study of personality in (cross-)cultural psychology is the cultural origin of the measurement instrument. Imported personality instruments are more likely to run into bias problems because they may be inadequate in tapping the underlying personality constructs outside their culture of origin (e.g., see Cheung & Leung, 1998). Depending on the cultural origin, three types of instruments (or studies) can be envisaged. The first can be called monocentered and utilizes an instrument of a single, usually western cultural background. Examples are found in the Big-Five and Eysenck traditions. A second approach uses decentered instruments. Researchers from various cultures jointly develop an instrument. A potential problem of the use of imported instruments under the monocentered approach is that important personality dimensions and processes that are obscure in the cultures in which the instruments are developed are unlikely to be discovered. This problem is much less severe when a decentered instrument is used, which will be valid and comprehensive all for cultures from which the instruments are derived. In a neighboring field, Schwartz's (1992) decentered approach in the development of a taxonomy of values is a good model for personality researchers to emulate. The approach is useful to identifying universally shared aspects of personality as well as culture-specific personality dimensions. The third type of instrument is multicentered. These studies are also known as cross-indigenous (Church, 1987) and convergent (Campbell, 1986). Ideally, it amounts to an experimental crossing of cultures and instruments. A researcher from each participating culture designs his or her own instrument, each of which will be administered in all cultures. The

approach can identify both universal and culture-specific aspects of personality, which show up as shared and country-specific sources of variance, respectively. Unfortunately, there are not many well-developed indigenous theories of personality, which impedes the development and administration of multicentered instruments (Berry et al., 1992; Sinha, 1997). In the sixth section we describe some examples.

Sampling of Cultures

Three types of schemes to sample cultures can be envisaged (Van de Vijver & Leung 1997a, b). The first is probability (or random) sampling. Because of the large cost of a probability sample from all existing cultures, the number of cultures from which samples are drawn may be restricted (e.g., to Western cultures) and a stratified (random) sampling may be carried out. The second and most frequently utilized type is convenience sampling. The choice of cultures is governed here by availability and cost efficiency. In the third type, called systematic sampling (also called purposive and judgmental sampling), the choice of cultures is more based on substantive considerations. It is useful when a personality theory is based on a particular taxonomy of cultures (e.g., individualism-collectivism and wealth) and a culture can be selected on the basis of its specific characteristics to provide an optimal test of the hypotheses. In generalizability studies, this approach often amounts to maximizing cultural differences (e.g., comparing a tight and a loose culture).

In personality research the by far most common type is convenience sampling. The instruments and theories that have been tested most extensively in (cross-)cultural psychology, such as the NEO-PI-R, EPQ,

MMPI-2, and the California Psychological Inventory (Gough, 1987) all focus on generalizability of findings: Do findings obtained in Euro-American countries also apply elsewhere? Convenience sampling has an obvious drawback as it usually leads to an overrepresentation of affluent countries, such as the USA, Western Europe, Korea, and Japan. As a consequence, the cross-cultural applicability of a theory or instrument can be easily overrated, as cultural diffusion, which is likely to be more extensive among affluent countries, tends to make these cultures similar. This bias in sampling may inflate the degree of cultural similarities that have been documented. For those who subscribe to a universal view of personality, it is definitely more informative to extend their work to poor, illiterate rural dwellers of some African countries or evaluate their instruments in culturally very diverse regions, such as the Middle East. Seen from this perspective, it is reassuring to see the variation of language families in which the MMPI (Butcher, Lim, & Nezami, 1998) and the NEO-PI-R (McCrae & Costa 1997; McCrae et al., 1998) have now been examined.

The focus on generalizability studies and convenience sampling highlights the paucity of theories that link personality and cultural dimensions. There are various typologies of cultures that could or are assumed to have a bearing on personality, such as field (in)dependence (that is assumed to influence agreeableness), tight versus loose cultures, and modern versus traditional cultures (Inkeles & Smith, 1974). MacDonald (1998) describes evolutionary approaches to the study of personality (see also Rushton, 1995). However, most work in (cross-)cultural psychology on personality revolves around individualism—collectivism (Hofstede, 1980; Triandis, 1995), and

studies in which predictions of cross-cultural differences and similarities are made on the basis of other cultural characteristics are hard to find. We are apparently in the stage of testing the universality of western theories and instruments; indigenous theories are not yet sufficiently developed to contribute to theories that link culture and personality.

Bias and Equivalence

Bias refers to the presence of nuisance or systematic error in a measure. For example, a poor item translation makes the score on the item (and strictly speaking, even the whole questionnaire) incomparable. Bias can be more formally defined as differences in scores between groups that have no correspondence in the domain of generalization of the test (e.g., Poortinga, 1989). Three types of bias can be envisaged (a more elaborate description can be found in Van de Vijver & Leung, 1997a, b; see also Van de Vijver, 2000). The first is construct bias; it occurs when the construct measured is not identical across cultures or when behaviors that constitute the domain of interest from which items are sampled are not identical across cultures. In Cheung et al.'s (1996) indigenous research on openness, the fifth factor of the five-factor model of personality did not emerge as a separate dimension of Chinese personality, which provides some evidence for the construct bias of this factor. Qualitative research is pivotal in the study of construct bias. Even if an imported instrument yields a structure identical to that found in the culture of origin of the instrument, it is still well possible that the instrument is “underinclusive” and does not cover all aspects relevant to the construct in the new culture.

The second is called method bias; it can result from sample incomparability, instrument characteristics, tester and interviewer effects, and the method (mode) of administration. Examples include (culturally different) social desirability and response sets. Method bias often leads to a change in mean score of a cultural group that can be easily misinterpreted as a valid cross-cultural difference, where it should be interpreted as a measurement artifact (such as some response style). The last type is item bias or differential item functioning (Camilli & Shepard, 1994; Holland & Wainer, 1993). It refers to anomalies at the item level, such as poor translations or inapplicability of an item in a specific culture.

Equivalence refers to the implications of bias on the comparability of constructs and test scores. Four types of outcomes of an examination of the comparability can be envisaged. The first type is labeled construct nonequivalence. It amounts to comparing “apples and oranges”. Because there is no shared attribute, no comparison can be made. The second is called structural (or functional) equivalence. An instrument administered in different cultural groups shows structural equivalence if the constructs it measures show similar internal structures (such as factor structures) and similar relationships with other variables. The third type is labeled measurement unit equivalence. This type of equivalence assumes interval- or ratio-level scores. Instruments show this type of equivalence if their measurement scales have the same unit of measurement and a different origin. If all items of an inventory are shifted upwards by the same amount in one cultural group (due to differential scale usage or social desirability), the measurement unit is identical across groups, but the means cannot be directly

compared due to the scale offset. Only in the case of scalar (or full score) equivalence direct comparisons can be made; it is the only type of equivalence that allows for statistical tests to compare means (such as t tests and analyses of variance).

A perusal of the literature in (cross-)cultural psychology shows that tests of equivalence and bias are not routinely applied. This is regrettable; without a test of equivalence it is impossible to know to what extent scores or constructs underlying an instrument can be compared across cultures.

Statistical Analysis of Bias and Equivalence

Structural Equivalence

The most commonly examined type of equivalence is structural equivalence. Two types of procedures have been extensively used to test for this type of equivalence. The first is exploratory factor analysis followed by target rotation and an evaluation of the factorial agreement across the samples (see, e.g., Van de Vijver & Leung, 1997b, for details). The most frequently employed agreement (or congruence) coefficient is Tucker's ϕ ; values larger than .90 are often taken to indicate equivalent factors. More recently, this value has been challenged by Van de Vijver and Poortinga (1994), who showed in a simulation study that values substantially higher than .90 can be obtained even when one or two items show markedly different loadings on factors with high eigenvalues (factors with many items with high loadings). The major problem with congruence coefficients is that statistical tests cannot be performed on these coefficients. To address this problem, Chan, Ho, Leung, Chan, and Yung (1999) propose to use a bootstrap procedure to estimate the standard error of Tucker's ϕ . A visual inspection

of differences in loadings of the target matrix and rotated source matrix may help to identify anomalous items.

Applications can be found in the work on the translation of the Big Five model of personality. McCrae, Costa, Del Pilar, Rolland, and Parker (1998) compared the equivalence of Filipino and French translations of the NEO-PI-R. All congruence coefficients were well above .90, in line with the earlier observed replicability of the factor structure of the NEO-PI-R across languages (McCrae & Costa, 1997). An example involving the Big Five model based on lexicons is provided by de Raad, Perugini, Hrebícková, and Szarota (1998). Other examples can be found in the cross-cultural work with the Eysenck Personality Questionnaire (e.g., Eysenck & Eysenck, 1983). The target rotations that were originally applied in this tradition have been criticized as insufficiently powerful to detect anomalous items (Bijnen, Van der Net, & Poortinga, 1986). More recently, an improved procedure has been proposed that yields the same results: The three Eysenck factors, psychoticism, extraversion, and neuroticism, are “strongly replicable” (Barrett, Petrides, Eysenck, & Eysenck, 1998, p. 805).

A less popular alternative to exploratory factor analysis is the usage of confirmatory factor analysis. Depending on whether the analysis is based on correlations or covariances, the analysis addresses structural or measurement unit equivalence, respectively. An example is provided by Windle, Iwawaki, and Lerner (1988), who investigated the cross-cultural stability of the Revised Dimensions of Temperament Survey among Japanese and American preschoolers. Using confirmatory factor analysis, evidence was found for the stability of loadings for each of nine factors (temperament

constructs) of the scale. Ghorpade, Hattrup, and Lackritz (1999) examined the equivalence of Rotter's (1966) Locus of Control scale, Rosenberg's (1965) Self-Esteem Scale, and some items of a need strength scale (Warr, Cook, & Wall, 1979) among participants from the USA and India. Separate confirmatory factor models were fit for men and women. Measurement unit equivalence could only be demonstrated for the Locus of Control Scale. The other scales yielded a poor fit in the group of Indian women, which was ascribed to their lower assimilation to western culture (as compared to Indian men).

That there are not many applications of confirmatory factor analysis in the establishment of structural equivalence in personality research may be due to various reasons. First, confirmatory techniques are less well known. Second, they are only appropriate when there is an a priori classification of items. Third, exploratory and confirmatory techniques do not always yield identical conclusions about structural equivalence; more specifically, the latter seems to provide a stricter test and may lead to a rejection of the hypothesis of structural equivalence for psychologically unclear or trivial reasons (e.g., (e.g., Katigbak, Church, & Akamine, 1996; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). Small and psychologically inconsequential differences in correlations across cultures often lead to a poor fit. Chan et al. (1999) suggested that target rotation with a bootstrap procedure for assessing factorial invariance statistically is a better procedure than confirmatory factor analysis.

Although the comparison of various factor analytic methods has only been carried out for the five-factor model, the results may be general. The

question arises then which technique should be preferred. Our proposal is that since there is no conclusive evidence for the superiority of any single approach, various approaches should be attempted. If the exploratory approach is used, the bootstrap procedure proposed by Chan and colleagues (1999) should be used so that a statistical evaluation of invariance rather than rules of thumb are possible. On the other hand, in applications of confirmatory techniques it is important to attempt to disentangle psychologically trivial model deviations from salient deviations. In particular, the overall chi square statistic may have limited usefulness. Model comparisons derived from a set of nested models may show which parameters are crucial to structural equivalence. Furthermore, factor loadings often constitute the “heart of the model” while factor correlations and error variances are often considered less important. Cross-cultural variation in the latter parameters may be less consequential for structural equivalence. Finally, it may be noted that in some applications different hypotheses are examined in exploratory and confirmatory analyses. In the latter, identity of factor loadings is often tested (in addition to equality of factor correlations and error variances), while in exploratory factor analyses the weaker hypothesis is tested that factor loadings are identical after correcting for differences in eigenvalues of the factors.

Other multivariate structure-oriented techniques such as multidimensional scaling and cluster analysis can also be utilized to examine structural equivalence. Unfortunately, these techniques are hardly ever used. In a recently defended Ph.D. study, Fontaine (1999) used multidimensional scaling to examine item and construct bias in Schwartz’s values instrument.

Structural equivalence can also be examined by studying the nomological network (external correlates) of an instrument. The technique is often applied when an instrument has been translated and no reference data from the original instrument are available. Examples are described by Paunonen and Ashton (1998).

It should be noted that even if a personality inventory shows structural equivalence in a (cross-)cultural study, the scores based on this instrument may not be comparable across cultures. In more technical terms, structural equivalence does not yet imply measurement unit equivalence, let alone full score equivalence. Procedures to establish scalar equivalence are mentioned only briefly here, as these are employed infrequently in (cross-)cultural studies of personality. Two kinds of approaches have been proposed. The first kind is direct and takes scalar equivalence to have been demonstrated if the assumptions of a usually stringent psychometric model have been met in the data. The best known examples are Item Response Theory and Structural Equation Modeling. If data in all cultural groups studied follow the same underlying model (as demonstrated by a nonsignificant fit statistic), scores are taken to be fully comparable and scores can be equally well compared within and across cultural groups. The only problem with this approach is the inability of all the statistical methods mentioned to distinguish between valid cross-cultural differences and those that are due to method bias. Therefore, Poortinga and Van de Vijver (1987) have proposed a less direct approach by arguing that additional measures are needed to validate a particular interpretation of score differences. For example, the measurement of social desirability can be used to rule out (or confirm) an alternative interpretation of

cross-cultural differences while the application of a multimethod approach addresses the important question to what extent an observed cross-cultural difference can also be found using another mode (Van de Vijver & Leung, 1997b, chapter 4).

Response Biases

It is known from cross-cultural research that social desirability, extreme responding, and acquiescence can produce artificial cultural differences, although correction for the biases does not always reduce these differences (Grimm & Church, 1999). We do not refer here to item-specific effects that would have been identified in an item bias analysis, but to more global influences on test scores. In a meta-analysis based on scores on Eysenck's Lie Scale, which measures social desirability, Van Hemert, Van de Vijver, Poortinga, and Georgas (2000b) found that there was a negative, significant correlation between the social desirability obtained in a country and the Gross National Product (GNP) (per head) of the country. As the latter is a proxy covering a wide range of variables ranging from education to health care to individualism, the mechanisms producing the correlation are not clear.

Marín, Gamba, and Marín (1992) found that Hispanics tended to show more extreme responses than did non-Hispanic whites and that this tendency disappeared with the level of adaptation of the Hispanics. Hui and Triandis (1989) found the same cultural difference when five-point rating scales were used, but reported that the tendency disappeared when ten-point scales were used. Stening and Everett (1984) found in a cross-cultural study among managers and expatriates from Britain, Hong Kong, Indonesia, Japan, Malaysia, the Philippines, Singapore, and Thailand that younger and better

educated respondents tended to use more midpoint responses. In a survey among American whites, Hispanics, and Mexicans, Ross and Mirowsky (1984) found more acquiescence among persons who were older, of Mexican origin, or lower in socioeconomic status. Van Herk (2000) examined acquiescence in an international marketing survey involving several European countries. She found that acquiescence was more prevalent among older and less educated respondents.

The convergence of findings of cross-cultural literature on response styles is noteworthy even though the small empirical basis precludes definite conclusions. The triplet of age, education, and socioeconomic status (that are often interrelated) tends to be related to response styles. Persons who are older, less educated, or come from lower socioeconomic strata are more likely to display response styles. Clearly, the likelihood of finding group differences due to response styles may increase with the cultural distance of the groups.

Not much is known about how these response biases affect cross-cultural comparison of personality data. If the bias is consistent across most items, such as an acquiescence bias or a nay-saying tendency, the effect should be minimal when structural equivalence is the key concern. If central tendency or extremity bias is uniform across items, again, the problem may not seriously threaten structural equivalence, because the factor loadings may not be sensitive to these biases.

Integrating Personality Instruments of Different Cultural Origins

Combining Qualitative and Quantitative Approaches in Developing Indigenous Instruments

In developing a personality instrument, the first step typically employs a qualitative approach for gaining insight into the content that should be covered by the instrument. The findings then provide the basis for developing the items contained in an objective personality instrument. For instance, the development of the MMPI is partly based on qualitative work conducted in an earlier stage (Butcher, Graham, Williams, & Ben-Porath, 1990; Williams, Butcher, Ben-Porath, & Graham, 1992). Unstructured interviews, open-ended questionnaires, the lexical approach (in which personality-related words are typically extracted from a dictionary of the language), and content analysis of written materials are common strategies that qualitative researchers use to capture the raw materials for developing a quantitative personality scale. The development of personality instruments in a non-Western context follows more or less the same path. For instance, some of the personality research conducted in the Philippines assumes an open-ended format that allows respondents to freely describe their views and feelings, upon which objective instruments are based (Church, 1987; Church & Katigbak, 1988, 1989). A similar approach of combining the qualitative and quantitative approaches was also followed in developing the Chinese Personality Assessment Inventory (CPAI) in Hong Kong and mainland China (Cheung et al., 1996). Extensive qualitative work, including interviews, content analysis of a variety of written materials, and open-ended questionnaires were used to identify descriptions for designing the objective personality instrument.

Joint Factor Analysis of Different Instruments

When several personality instruments derived from different cultural contexts are available, it becomes interesting to compare them. The problem

assumes the form of a multisample multi-instrument matrix. The simplest and most common case is a 2×2 matrix, with two cultures, in each of which an instrument is being developed. The ideal design calls for administering the two instruments to both cultural groups. In reality, an American inventory is often administered, together with a locally developed instrument (e.g., Katigbak, Church, & Akamine, 1996; Yang & Bond, 1990; Yik & Bond, 1993). An exploratory factor analysis is then performed on these two instruments, the indigenous and the American instruments. Typically, the factor structures are examined in terms of (1) whether the hypothesized factor structure of the American instrument was confirmed in this culture, and (2) whether the factors of the indigenous instruments are similar to the factors derived from the American instrument. The typical methods for determining the replicability of the American factors are well-known and have been described before. In contrast, there is still no widely accepted way to determine the overlap of imported and indigenous instruments. The common practice is to try a number of solutions with different number of factors, and researchers then evaluate which solution is optimal. The overlap of American and indigenous factors are determined by subjective judgment (e.g., Yang & Bond, 1990). Recently, in determining the overlap between the five factor model represented by the NEO-PI-R and the CPAI, in addition to judging the optimality of various factor solutions, Cheung et al. (in press) also used regression to evaluate whether a factor that is defined by CPAI scales only, but not by NEO-PI-R facets can be subsumed by the Big Five model. Each of the four scales defining this new factor, interpersonal relatedness, was regressed on the five personality factors measured by NEO-PI-R in two different samples. The resulting R^2

ranged from .08 to .31, with a mean of .17. Based on this relatively small value, Cheung et al. argued that the interpersonal relatedness factor is distinct from the Big Five factors.

We propose that internal analysis alone, such as deciding on the optimal number of factors, may not be adequate to address the issue of whether distinct personality factors missed by imported instruments are being identified by indigenous instruments. The use of external correlates provides another avenue to address this issue. If the factors claimed to be distinct can generate additional variance of some external variables over and beyond those factors that are claimed to be all-encompassing, their distinctiveness is hard to be explained away. For examples of this approach, see Yik and Bond (1993) and Zhang and Bond (1998).

Personality as an Individual and as a Cultural Characteristic

Historically, there are two lines of theory and research that have attempted to study individual and cultural characteristics and their relationships. The first comes from anthropology and is known as culture-and-personality and more recently, psychological anthropology (Bock, 1980; Piker, 1998). Much of this work, which is often based on the applications of psychodynamic concepts to cultural institutions, did not receive warm support in (cross-)cultural psychology. For example, Berry et al. (1992) argue that by focusing on psychological aspects shared by all members of a culture these studies tend to underemphasize individual differences; in more statistical terms, the anthropologically based approaches of personality overemphasize the between-culture variance and underemphasize the within-culture variance. We know from the literature on cross-level inferences that

the application of results obtained at one level to another level is fraught with problems (Achen & Shively, 1995). For example, that 2% of the women in a country are pregnant does not imply that any individual woman is 2% pregnant (cf. Robinson, 1950). The second line comes from large-scale multicultural studies of psychological attributes, such as Hofstede's (1980) model of country differences. The same lack of attention for within-culture individual differences remains. In conclusion, both lines of theory and research focused on country-level psychological characteristics and failed to scrutinize the link between individual and country characteristics, addressing questions such as the identity of characteristics at individual and country level (Is uncertainty avoidance the same at both levels?) and their interaction (e.g., what is the difference between being an collectivist in an individualist and a collectivist country?). Neither the anthropological nor the psychological approach was adequately capturing the relationship between individuals and culture.

The development of multilevel models (Bryk & Raudenbush, 1992; Goldstein, 1987; Muthén, 1991, 1994) may give a new impetus to the modeling of individual—culture relationships (Van de Vijver & Leung, 2000). Two types of multilevel models are relevant for personality—culture studies. The first type is level oriented and assumes that variables have been measured at individual and cultural levels (Bryk & Raudenbush, 1992; Goldstein, 1987). For example, Maas and Meijnen (1999) studied factors that influence whether or not teachers define a child as having learning or behavior problems. Data on 2,000 fourth-graders from 114 primary school classes in the Netherlands were examined. Variation in teacher ratings were

studied at individual and class levels; 87% of the variation was related to the individual level while merely 13% was class related.

The second and for personality more important type of multilevel models are structure oriented. They aim at comparing the structure of a construct at individual and cultural levels. Muthén (1991, 1994) has developed a procedure for multilevel confirmatory factor analysis, whereas Van de Vijver and Poortinga (2000) developed a procedure for exploratory factor analysis. The procedure by Van de Vijver and Poortinga amounts to a comparison of the structures at individual and cultural levels, somewhat similar to the procedure described before to compare factor structures obtained in different cultures. In the first step, two factor analyses are carried out: one based on the pooled data (the covariance matrices of the separate countries are averaged to form a single pooled-within covariance matrix), and the other one based on the country-level data (each country constitutes one observation). In the second step, one set of factor loadings (either individual or country level) are rotated to the other set, and factor congruence is evaluated. If (and only if) the agreement is high, it is concluded that the constructs have the same meaning at individual and cultural levels.

An example of this approach is provided by Van Hemert et al. (2000b), who performed a secondary analysis of the data obtained with the Eysenck Personality Questionnaire from 24 countries. At neither level were the scales independent of each other; furthermore, the between-country correlations were stronger than the within-country correlations (due to the reduction of random error by aggregation). A multilevel factor analysis showed that equality of meaning at individual and country levels could be established for

the extraversion and neuroticism factor only. As discussed earlier, the Lie Scale showed a strong and significant correlation of $-.67$ with GNP. No other scale showed a significant correlation with GNP (psychoticism: $-.19$; extraversion: $.11$; neuroticism: $-.06$). Aggregating individual lie scale scores at country level introduces a source of variation that is not shared by the other scales. The lack of equivalence of the Psychoticism Scale may be due to measurement problems of the scale reported before (e.g., Goh, King, & King, 1982); the internal consistency of the scale and score variability are not high, mainly due to the many scores piling up at the low end of the scale. On the other hand, score aggregation may have changed the meaning of the scale. Whereas psychoticism and religiosity are related at individual level, no such relationship was found at country level. Our analyses strongly suggest that there is no structural equivalence of all EPQ scales at individual and country level.

Multilevel analysis often involves external linkage analyses as well. In the Van Hemert et al. (2000b) study, the four personality scales were found to be related to some country-level variables. Death from political violence showed a significant, positive correlation with Extraversion ($.35$). Psychoticism was negatively related to subjective well-being and religiosity ($-.41$ and $-.35$, respectively), and positively to Hofstede's masculinity ($.47$). Neuroticism showed a significant, positive correlation with Masculinity ($.63$). The most striking finding was the significant, negative correlation between the Lie Scale and Gross National Product (and related economics variables).

Country-level analyses such as carried out here have some difficulties. The sample size was small, as each country constituted a single case.

Economical indicators are available for most countries, but the social indicators are available only for a small set of western countries. Some correlations between scale scores and other variables were similar at individual and country level. For instance, at country level extraversion showed a positive correlation with subjective well-being, which replicates findings at individual level. That the correlation did not reach significance at country level may be due to sample size. Furthermore, cultural variation in social indicators is often small, and cultural diffusion becomes severe as western countries have mutual influence on each other. The correlations found may be a poor estimate of the population values. There seems to be no easy solution for these difficulties.

Future Directions

There are two themes in the study of methodological aspects of personality research in (cross-)cultural psychology that require further examination. The first one deals with the robustness of procedures to establish equivalence. Two recent studies have examined equivalence in great detail (Byrne & Campbell, 1999; Chan et al., 1999). Both found that even when there is evidence of structural equivalence based on commonly accepted standards, noticeable differences between subgroups of participants or between items (when data are analyzed at the scale level) might still exist. These results seem to imply that, like assumptions of statistical techniques, conditions for equivalence are almost never fully met in empirical data. Future equivalence research should focus on implications of distortions of equivalence conditions on the comparability of constructs and scores, analogous to the studies of robustness of statistical techniques. For example,

suppose that a researcher finds cross-cultural differences in loadings of a scale, leading to a value of Tucker's phi well below .90. This finding could point to either of two psychologically very different outcomes. First, the differences in loadings may be due to poor item translations and inadequate item contents for some countries; however, such differences would probably be seen as largely technical, which after suitable modification of the instrument or elimination of the poor items would not challenge the structural equivalence. Second, the differences in loadings may be found on conceptually related items in which case it would have to be concluded that the measure shows structural equivalence for only part of the construct, possibly leading to a redefinition of the original concept (e.g., introducing a specification of the domain to which the concept applies). Third, item elimination may hardly change the agreement index; in such a case the lack of equivalence is psychologically more consequential and the underlying constructs show limited comparability, presumably making it necessary to use different labels for the construct in different cultures. In sum, there is a need to further specify which distortions of equivalence are relevant and which distortions can be ignored for certain applications.

Another line of research is related to the currently weak link between individual and culture. It has turned out to be a difficult task to link individual- and culture-level characteristics. Personality psychology has found it difficult to apply its concepts to culture-level phenomena; analogously, concepts and theories of psychological anthropology could not be fruitfully applied at the individual level. Seen from this perspective, personality and culture display an unhealthy Cartesian dualism. Bridging the gap between personality and

culture is obviously not only a methodological issue; it also requires new theories and more interest in the joint study of individual and cultural characteristics. The usually strong intra-disciplinary focus of psychologists is not conducive to such an interdisciplinary undertaking. Yet, multilevel models may provide a useful new tool for this endeavor.

References

- Achen, C. H., & Shively, W. P. (1995). Cross-level inference. Chicago: University of Chicago Press.
- Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (1998). The Eysenck Personality Questionnaire: An examination of the factorial similarity of P, E, N, and L across 34 countries. Personality and Individual Differences, *25*, 805-819.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R., (1992). Cross-cultural psychology. Research and applications. Cambridge: Cambridge University Press.
- Bijnen, E. J., Van der Net, T. Z., & Poortinga, Y. H. (1986). On cross-cultural comparative studies with the Eysenck Personality Questionnaire. Journal of Cross-Cultural Psychology, *17*, 3-16.
- Bock, P. K. (1998). Continuities in psychological anthropology. San Francisco: Freeman.
- Boesch, E. (1996). The seven flaws of cross-cultural psychology: The story of a conversion. Mind, Culture, and Activity, *1*(3), 2-10.
- Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis. Newbury Park, CA: Sage.
- Butcher, J. N., Graham, J. R., Williams, C. L., Ben-Porath, Y. S. (1990). Development and use of the MMPI-2 content scales. Minneapolis: University of Minnesota Press.
- Butcher, J. N., Lim, J., & Nezami, E. (1998). Objective study of abnormal personality in cross-cultural settings. The Minnesota Multiphasic

Personality Inventory (MMPI-2). Journal of Cross-Cultural Psychology, 29, 189-211.

Byrne, B. M., & Campbell, T. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. Journal of Cross-Cultural Psychology, 30, 555-574.

Camilli, G., & Shepard, L. N. (1994). Methods for identifying biased test items. Thousand Oaks, CA: Sage.

Campbell, D. T. (1986). Science's social system of validity-enhancing collective believe change and the problems of the social sciences. In D. W. Fiske & R. A. Shweder (Eds.), Metatheory in social science (pp. 108-135). Chicago: University of Chicago Press.

Chan, W., Ho, R. M., Leung, K., Chan, D. K-S., & Yung, Y-F. (1999). An alternative method for evaluating congruence coefficients with Procrustes rotation: A bootstrap procedure. Psychological Methods, 4, 378-402.

Cheung, F. M., & Leung, K. (1998). Indigenous personality measures. Chinese examples. Journal of Cross-Cultural Psychology, 29, 233-248.

Cheung, F. M., Leung, K., Fan, R. M., Song, W. Z., Zhang, J. X., & Chang, J. P. (1996). Development of the Chinese Personality Assessment Inventory. 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. (in press). Indigenous Chinese personality constructs: Is the five factor model complete? Journal of Cross-Cultural Psychology.

Church, A. T. (1987). Personality research in a non-Western culture: The Philippines. Psychological Bulletin, 102, 272-292.

Church, A. T., & Lonner, W. J. (1998). The cross-cultural perspective in the study of personality: Rationale and current research. Journal of Cross-Cultural Psychology, 29, 32-62.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

de Raad, B., Perugini, M., Hrebícková, M., & Szarota, P. (1998). Lingua franca of personality: Taxonomies and structures based on the psycholexical approach. Journal of Cross-Cultural Psychology, 29, 212-232.

Dunbar, E., Saiz, J. L., Stela, K., & Saez, R. (2000). Personality and social group value determinants of out-group bias. A cross-national comparison of Gough's Pr/To Scale. Journal of Cross-Cultural Psychology, 31, 267-275.

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.

Fiske, A. P. (1991). Structures of social life: The four elementary forms of human relations : communal sharing, authority ranking, equality matching, market pricing. New York: The Free Press.

Fontaine, J. (1999). Culturele vertekening in Schwartz' waardeninstrument: een exemplarisch onderzoek naar culturele vertekening in sociaal-psychologische en persoonlijkheidsvragenlijsten [Cultural bias in Schwartz's values instrument: A case study of the cultural bias in social psychological and personality inventories]. Ph.D. Dissertation, University of Leuven, Leuven.

Ghorpade, J. Hattrup, K. & Lackritz, J. R. (1999). The use of personality measures in cross-cultural research: A test of three personality scales across two countries. Journal of Applied Psychology, 84, 670-679.

Goh, D. S., King, D. W., & King, L. (1982). A psychometric evaluation of the Eysenck Personality Questionnaire. Educational and Psychological Measurement, 42, 297-309.

Goldstein, H. (1987). Multilevel models in educational and social research. London: Griffin.

Gough, H. G. (1987). Administrator's guide for the California Psychological Inventory. Palo Alto, CA: Consulting Psychologists Press.

Greenfield, P. M. (1997a). Culture as process: Empirical methods for cultural psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), Handbook of cross-cultural psychology (2nd ed., vol. 1, pp. 301-346). Boston: Allyn & Bacon.

Grimm, S. D., & Church, A. T. (1999). A cross-cultural study of response biases in personality measures. Journal of Research in Personality, 33, 415-441.

Holland, P. W., & Wainer, H. (Eds.) (1993). Differential item functioning. Hillsdale, NJ: Erlbaum.

Hsu, F. K. (1970). Americans and Chinese: Two ways of life. New York: Doubleday.

Hui, C. H., Triandis, H. C. (1983). Multistrategy approach to cross-cultural research: The case of locus of control. Journal of Cross Cultural Psychology, 14, 65-83.

Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. Journal of Cross-Cultural Psychology, 20, 296-309.

Inkeles, A., & Smith, D. H. (1974). Becoming modern. Cambridge, MA: Harvard University Press.

Jahoda, G. (1995). In pursuit of the emic-etic distinction: Can we ever capture it? In N. R. Goldberger & J. B. Veroff (Eds.), The culture and psychology reader (pp. 128-138). New York: New York University Press.

Katigbak, M. S., Church, A. T., & Akamine, T. X. (1996). Cross-cultural generalizability of personality dimensions: Relating indigenous and imported dimensions in two cultures. Journal of Personality and Social Psychology, 70, 99-114.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations for life satisfaction: Adding relationship harmony to self-esteem. Journal of Personality and Social Psychology, 73, 1038-1051.

Lonner, W. J. (1980). The search for psychological universals. In H. C. Triandis & W. W. Lambert (Eds.), Handbook of cross-cultural psychology (Vol. 1, pp. 143-204). Boston: Allyn & Bacon.

Maas, C. J. M., & Meijnen, G.W. (1999). Problem students: A contextual phenomenon? Social Behavior and Personality, 27, 387-406.

MacDonald, K. (1998). Evolution, culture, and the five-factor model. Journal of Cross-Cultural Psychology, 29, 119-149.

Marín, G., Gamba, R. J., & Marín, B. V. (1992). Extreme response style and acquiescence among Hispanics: The role of acculturation and education. Journal of Cross-Cultural Psychology, 23, 498-509.

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

McCrae, R. R., Costa, P. T., Del Pilar, G. H., Rolland, J-P., & Parker, W. D. (1998). Cross-cultural assessment of the five factor model. The Revised NEO Personality Inventory. Journal of Cross-Cultural Psychology, 29, 171-188.

McCrae, R. R., Yik, M. S. M., Trapnell, P. D., Bond, M. H., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. Journal of Personality and Social Psychology, 74, 1041-1055.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. Journal of Personality and Social Psychology, 70, 522-566.

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.

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.

Morris, M. W., Leung, K., Ames, D., & Lickel, B. (1999). Views from inside and outside: Integrating emic and etic insights about culture and justice judgment. Academy of Management Review, 24, 781-796.

- Naroll, R., & Cohen, R. (Eds.), A handbook of cultural anthropology. New York: American Museum of Natural History.
- Paunonen, S., & Ashton, M. C. (1998). The structured assessment of personality across cultures. Journal of Cross-Cultural Psychology, 29, 150-170.
- Piker, S. (1998). Contributions of psychological anthropology. Journal of Cross-Cultural Psychology, 29, 9-31.
- Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. International Journal of Psychology, 24, 737-756.
- Poortinga, Y.H. (1997b). Towards convergence? In J. W. Berry, Y. H. Poortinga & J. Pandey (Eds.), Handbook of cross-cultural psychology (2nd ed., Vol.1, pp. 301-346). Boston: Allyn & Bacon.
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. American Sociological Review, 15, 351-357.
- Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.
- Ross, C. E., & Mirowsky, J. (1984). Socially-desirable response and acquiescence in a cross-cultural survey of mental health. Journal of Health and Social Behavior, 25, 189-197.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80(1), 1-28.
- Rubinstein, G. (1996). Two peoples in one land: A validation study of Altemeyer's Right-Wing Authoritarian scale in the Palestinian and Jewish societies in Israel. Journal of Cross-Cultural Psychology, 27, 216-230.

Rushton, J. P. (1995). Race, evolution, and behavior. New Brunswick, NJ: Transaction.

Sinha, D. (1997). Indigenizing psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), Handbook of cross-cultural psychology (2nd ed., vol. 1, pp. 131-169). Boston: Allyn & Bacon.

Stening, B. W., & Everett, J. E. (1984) Response styles in a cross-cultural managerial study. Journal of Social Psychology, 122, 151-156.

Triandis, H. C. (1995). Individualism and collectivism. Boulder, CO: Westview.

Van de Vijver, F. J. R. (2000). Research methods. D. Matsumoto (Ed.), Handbook of culture and psychology. Oxford: Oxford University Press. (in press)

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., & Leung, K. (2000). Methodological issues in psychological research on culture. Journal of Cross-Cultural Psychology, 31, 33-51.

Van de Vijver, F. J. R., & Poortinga, Y. H. (1994). Methodological issues in cross-cultural studies on parental rearing behavior and psychopathology. In C. Perris, W. A. Arrindell, & M. Eisemann (Eds.), Parental rearing and psychopathology (pp. 173-197). Chichester: Wiley.

Van de Vijver, F. J. R., & Poortinga, Y. H. (2000). Structural equivalence in multilevel research. Tilburg University (in preparation).

Van Hemert, D. A., Van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2000a). The Beck Depression Inventory in 30 countries: A meta-analysis. Tilburg: Tilburg University. (in preparation)

Van Hemert, D. A., Van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2000b). Structure and score levels of the Eysenck Personality Questionnaire across individuals and countries. Tilburg: Tilburg University. (in preparation)

Van Herk, H. (2000). Equivalence in a cross-cultural context: Methodological & empirical issues in marketing research (PhD thesis). Tilburg: Tilburg university.

Warr, P., Cook, J., & Wall, T. (1979). Scales for the measurement of some work attitudes and aspects of psychological well-being. Journal of Occupational Psychology, 52, 129-148.

Williams, C. L., Butcher, J. N., Ben-Porath, Y. S., & Graham, J. R. (1992). MMPI-A content scales: Assessing psychopathology in adolescents. Minneapolis, University of Minnesota Press.

Windle, M., Iwawaki, S., & Lerner, R. M. (1988). Cross-cultural comparability of temperament among Japanese and American preschool children. International Journal of Psychology, 23, 547-567.

Yang, K. S., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. Journal of Personality and Social Psychology, 58, 1087-1095.

Yik, M. S., & Bond, M. H. (1993). Exploring the dimensions of Chinese person perception with indigenous and imported constructs: Creating a culturally balanced scale. International Journal of Psychology, 28, 75-95.

Zhang, J. X., & Bond, M. H. (1998). Personality and filial piety among college students in two Chinese societies: The added value of indigenous constructs. Journal of Cross Cultural Psychology, 29, 402-417.

Table 1 Types of Studies in (Cross-)Cultural Psychology (After Van de Vijver & Leung, 1997a, b)

Consideration of contextual factors	Orientation more on	
	Hypothesis testing	Exploration
No	Generalizability	Psychological differences
Yes	Contextual theory	Ecological linkage