The Between and Within of Cross-Cultural Research

Richard A. Shweder

INTRODUCTION

Cultures, and individuals within any culture, may not differ from one another in the same ways. This paper presents an analysis of data that indicate a discrepancy between across unit and within unit findings and discusses the possibility that valid cross-cultural variables are intraculturally inappropriate.

The lack of necessary correspondence between cultural and individual levels of analysis has been discussed by Minturn and Lambert (1964:255–257) from the point of view of the intracultural validation of cross-cultural findings. In this paper the question of noncorrespondence between levels of analysis is initially focused on

Richard A. Shweder is Assistant Professor of Human Development in the Committee on Human Development at the University of Chicago. He received his B.A. in anthropology from the University of Pittsburgh in 1966 and his Ph.D. in the Department of Social Relations from Harvard University in 1972. He has been a research fellow in the Department of Psychology and Social Relations at Harvard University, a fellow of the Child Development Research Unit at the University of Nairobi, and a lecturer in the Department of Sociology at the University of Nairobi. He has carried out field work in Mexico, India, and Africa.

This paper is a modification and expansion of one section of a manuscript ("Is a Culture a Situation?") that received honorable mention in the 1971 competition for the Stirling Award in Culture and Personality Studies. John Whiting, Beatrice Whiting, Thomas Weisner, and Wendy Jackson are to be thanked for the exchange of ideas that inspired that manuscript.
differences in the cross-cultural and intracultural structure of children's social behavior. Data on the structure of Rorschach test response categories and social behavior in small groups is subsequently discussed.

Whiting and Whiting (1972) have recently completed an analysis of the six culture data.\(^1\) Two theoretical dimensions of social behavior, nurturance versus egoism and intimacy versus aggression, are related to cultural complexity and family nucleation. Children are more egoistic (and less nurturant) in complex cultures, and more intimate (and less aggressive) in cultures where the nuclear family predominates. The American children in the Whitings’ sample are the most egoistic and the most intimate.

The point of departure for this paper is a problem I encountered trying to select a set of behaviors as indicators of intracultural variations in nurturance versus egoism. Using some intracultural data published by Longabaugh (1966:108–110) I discovered that the nurturance versus egoism dimension is valid at the cross-cultural level yet can not be applied intraculturally.

The operational indicators that are valid for classifying the 134 children from the six culture study as nurturant versus egoistic at the cross-cultural level are not valid within any of these six cultures. The variable can be used to describe cultural differences but not individual differences within a culture.

This lack of intracultural applicability of a valid cross-cultural variable in no way detracts from the importance of the Whitings’ cross-cultural findings. Instead it raises questions about the relationship between cross-cultural and intracultural variables and their indicators:

1. What are the criteria for assessing the validity of operational indicators?
2. Do cultural communities and individuals within these communities differ in the same ways?

\(^1\) The six culture study was an investigation of child rearing practices and childhood social behavior carried out in 1954–1955 under the direction of John Whiting, Irvin Child, and William Lambert (Minturn and Lambert 1964, Whiting 1965, Whiting and Whiting 1972). One aspect of the project was the observation and recording of the social behavior of 154 boys and girls of age 3 to 11 from six cultures. The data included more than 2,000 five-minute observations of social behavior in natural settings, and more than 10,000 social interactions. It was later coded in terms of a common set of twelve categories.
3. How is it possible for a valid cross-unit variable to lack appropriateness within the units?

**NURTURENCE VERSUS EGOISM ACROSS AND WITHIN CULTURES**

**THE CROSS-CULTURAL VARIABLE**

Nurturance versus egoism is a theoretical variable that distinguishes behaviors beneficial to others from behaviors beneficial to oneself. In the Whitings' analysis the three most valid operational indicators of nurturance are *offers help, offers support and approval,* and *makes responsible suggestions.* The three most valid operational indicators of egoism are *seeks attention, seeks help,* and *dominates.*

In table 1 the correlations (Pearson r) *across* six cultures among the six nurturance versus egoism indicators are presented. These correlations appeared in an early provisional manuscript by the Whitings and were reproduced in Shweder (1971:35). The correlations were calculated over the mean proportion scores for each type of behavior for each culture.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NURTURANCE VERSUS EGOISM: VALID CROSS-CULTURAL INDICATORS</strong></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1 Offers support and approval</td>
</tr>
<tr>
<td>2 Offers help</td>
</tr>
<tr>
<td>3 Makes responsible suggestions</td>
</tr>
<tr>
<td>4 Seeks attention</td>
</tr>
<tr>
<td>5 Dominates</td>
</tr>
<tr>
<td>6 Seeks help</td>
</tr>
<tr>
<td><strong>N = 6.</strong></td>
</tr>
</tbody>
</table>

The *pattern* of correlations in table 1 is compatible with the theoretical existence of nurturance and egoism as opposed features of social behavior. The three nurturance indicators are positively correlated among themselves. The three egoism indicators are positively correlated among themselves. The nurturance indicators and the egoism indicators are negatively correlated with each other.

**THE VALIDITY OF INDICATORS OF THEORETICAL VARIABLES**

Three validity criteria can be applied to the indicators in table 1. All three support the validity of these indicators as measures of nurturance versus egoism *at the cross-cultural level.* The three criteria are as follows:
1. **Meaningfulness.** Each indicator possesses a semantic attribute that corresponds to the theoretical definition of the variable it is supposed to measure. For example, offering help is a kind of nurturance (a benefiting of others) at the purely conceptual level.

2. **Internal Consistency.** Indicators that are meaningfully related to the same theoretical variable are empirically associated with one another. For example, the three nurturance indicators are positively correlated.

3. **External Consistency.** Each indicator of a theoretical variable \( V_1 \) is related to each indicator of another relevant theoretical variable \( V_2 \) in the same way. For example, each nurturance indicator is negatively correlated with each egoism indicator.

Semantic similarity is an obvious validity criterion. Operational indicators with no semantic connection to their common theoretical variable would lack "face validity" no matter how internally or externally consistent they might be. Internal consistency is the validity criterion predicted by the abstract representation of the causal relationship between theoretical variables and their indicators.

In this representation operational indicators are assumed to be *causally* related to their theoretical variables in a special and simplified way. The causal assumption is that a set of indicators are related to one another only as effects of the same influence, the theoretical variable. None of the indicators is assumed to be the direct cause or effect of another.

This simplifying assumption about the relationship of indicators to their theoretical variable means that covariations among indicators only occur because of variations in their common theoretical variable.

On the basis of this assumption certain predictions can be made about the way indicators of the same theoretical variable should be empirically related. There are two such predictions (see Blalock 1964:167-169):

1. All pairs of indicators of the same theoretical variable will have spurious positive correlations. The positive correlations are spurious because they reflect an indirect relationship among the indicators, i.e., they are only related to one another as effects of the same cause.

2. The positive correlations among all pairs of indicators will reduce to zero when the influence of the theoretical variable is statistically or experimentally controlled.
Internal consistency meets the first of these requirements. The internal consistency criterion eliminates theoretical variables that have no empirical existence or are poorly conceptualized; for example, those that are too global and undiscriminating.

External consistency refers to the mutual substitutability of operational indicators that are valid measures of the same theoretical variable. That is to say, the various indicators of variable \( V_1 \) and the various indicators of variable \( V_2 \) must give the same answer to the question “how is theoretical variable \( V_1 \) related to theoretical variable \( V_2 \) ?” (see Sears 1961:447). If two postulated indicators of the same theoretical variable are related to another variable in different ways they are not measuring the same thing.

At the cross-cultural level the indicators of nurturance and egoism conform to these three validity criteria (see table 1). Had they failed to meet any of these criteria serious questions would arise concerning their validity as measures of nurturance and egoism.

NURTURANCE VERSUS EGOISM: THE CROSS-CULTURAL AND INTRACULTURAL DISCREPANCY

One of the important cross-cultural findings of the Whitings (1972) is the positive relationship between cultural simplicity and the development of nurturant versus egoistic personalities in children. Simplifying somewhat, their causal argument runs as follows. Women from comparatively simple cultures (where occupational specialization is underdeveloped) have comparatively heavy work loads and thus frequently assign baby-sitting tasks to their children. Babies characteristically elicit nurturant behaviors from their child tenders. In proportion to the amount of their baby-sitting experience the child tenders behave nurturantly in non-baby-sitting interactions in comparison with children from other cultures.

The Whitings also ask whether the relationship between baby-sitting experience and nurturance can be discovered intraculturally. Yet to ask this it is necessary to locate a valid set of operational indicators to classify children within a single culture as more or less nurturant. Intraculturally the six cross-cultural indicators (see table 1) are not valid. This can be seen in tables 2 and 3.

Table 2 shows the correlations (Pearson \( r \)) among the six nurturance versus egoism indicators within the cultural community called Khalapur. The correlations from Longabaugh (1966:110) are based
Table 2

Nurturance Versus Egoism: Invalid Intracultural Indicators

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Offers support and approval</td>
<td>.17</td>
<td>-.25</td>
<td>.10</td>
<td>-.04</td>
<td>-.33</td>
</tr>
<tr>
<td>2 Offers help</td>
<td>.11</td>
<td>-.31</td>
<td>-.52</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>3 Makes responsible suggestions</td>
<td>-.12</td>
<td>-.12</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Seeks attention</td>
<td>.28</td>
<td></td>
<td>-.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Dominates</td>
<td></td>
<td></td>
<td>-.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Seeks help</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 24.

on the average proportion of each indicator in the behavior profiles of twenty-four Rajput children.

The six indicators in table 2 lack internal and external consistency at the intracultural level. Seeks help, and dominates (internally consistent indicators of egoism at the cross-cultural level) are negatively correlated (r = -.56). Offers help, and offers support and approval (both valid indicators of nurturance at the cross-cultural level) are related to seeks help (a valid indicator of egoism at the cross-cultural level) in different ways (r = +.14 and -.33 respectively).

It is apparent that the pattern of correlations in table 2 is no longer organized along a nurturance versus egoism dimension.

Either the six indicators are invalid measures of nurturance versus egoism within Rajput culture or the dimension itself is not appropriate to describe individual differences among these children.

In Table 3 I have listed the results of a comparison between the pattern of correlation coefficients among the six cross-culturally

Table 3

Nurturance Versus Egoism: The Intracultural Invalidity of Valid Cross-Cultural Indicators

<table>
<thead>
<tr>
<th></th>
<th>Degree of similarity of the pattern of indicator intercorrelations in each of six cultures to the valid pattern of intercorrelations among the indicators in table 1 (the cross-cultural pattern).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taira</td>
<td>-.40</td>
</tr>
<tr>
<td>Orchard Town</td>
<td>-.22</td>
</tr>
<tr>
<td>Khalapur</td>
<td>.10</td>
</tr>
<tr>
<td>Tarong</td>
<td>.12</td>
</tr>
<tr>
<td>Juxtlahuaca</td>
<td>.29</td>
</tr>
<tr>
<td>Nyansongo</td>
<td>.38</td>
</tr>
<tr>
<td>Average r,</td>
<td>.05</td>
</tr>
</tbody>
</table>

N = 15 cells of a six by six half matrix.

2. Data from Longabaugh (1966:108–110) are analyzed by Shweder.
valid indicators of nurturance versus egoism (table 1) and the pattern of correlations among these same indicators within each of six cultures.

The comparison was performed by ranking the cells of the matrix in table 1 in order of their magnitude and direction. This ranked list of matrix cells was compared separately with the ranked cells of each of the within culture matrices. In each of the six comparisons the two ranked lists were intercorrelated using the Spearman coefficient \( r_s \) as a measure of similarity.

The cross-cultural patterning of the six indicators (table 1) is quite different from the intracultural patterning of these indicators within each of the six cultures. In table 3 the degree of similarity of the cross-cultural matrix to each of the intracultural matrices ranges from \(-.40\) (Taira) to \(+.38\) (Nyangongo).

The average degree of similarity of these six intracultural matrices to the cross-cultural matrix is only \(.05\) \( r_s \). That is to say, the six indicators that are valid measures of nurturance versus egoism at the cultural level are not valid at the individual level within these cultures.

Of course, one might argue that the discrepancy between the intracultural and cross-cultural correlations results from unreliable data, the consequence of an insufficient number of observations on the small number of individuals observed within each culture (from sixteen to twenty-four). I doubt the discrepancy between the intracultural and cross-cultural patterning of indicator correlations can be dismissed in this way, for two reasons: (1) the cross-culturally valid indicators of the Whitings' second dimension—intimacy versus aggression—tend toward validity at the intracultural level, and (2) the discrepancy discovered between within and across unit correlations can be found in other types of data.

THE BETWEEN AND WITHIN UNIT DISCREPANCY: FURTHER EXAMPLES

I have found instances of a discrepancy between across and within unit correlations among Rorschach test response categories, and among Bales's interaction process analysis (IPA) categories.

THE RORSCHACH TEST

The Rorschach test has been administered to people from a vari-
ety of cultures. Ten standard inkblots have been shown to people all over the world, and their reactions to them have been coded in terms of a standard set of response categories.

I have analyzed "raw data" on individual responses to the Rorschach test for each of nine cultures. These references are stated in table 4.\(^3\) Several problems arose in the course of data analysis. The studies did not publish data on exactly the same set of response categories. Many of the response categories occurred too infrequently to warrant analysis by themselves. Many subjects had very low total response levels. To compensate for these limitations, a small common subset of response categories (some of which were combined) was selected. Supplementary data on the modal responses of a culture were used in the calculation of cross-cultural correlations; and some subjects were entirely dropped from the analysis.

Does the patterning of correlations among Rorschach response categories (e.g., the use of the whole inkblot \([W]\), the attribution of movement \([M]\), the use of pure form \([F]\)) across cultures correspond to the patterning of these same responses within any or all cultures?

I answered this question in the following way. The mean proportion scores for six Rorschach response categories \((W, M, F, A, FC + CF + C, KK + FK + Fc)\) were intercorrelated over nine cultures. These lettered variables respectively stand for the use of whole figures, movement, pure form, animal imagery, shading, and color. The matrix of correlation coefficients \((Pearson r)\) across these cultures was compared with an intracultural matrix among these same response categories within each of seven cultures. The cell-ranking procedure used for comparing the cross-cultural matrix with each of the intracultural matrices has been discussed above. The results of the comparison are shown in table 4.

Table 4 indicates that the intracultural patterning of Rorschach responses is usually different from the cross-cultural patterning of those same responses. Only one intracultural matrix (Mormon) is similar to the cross-cultural matrix. The average degree of similarity between the intracultural matrices and the cross-cultural matrix is only \(0.19 (r_s)\).

In table 4 the discrepancy between within and across unit correlations appears again. If the finding applies to other Rorschach response categories as well, a factor analysis of the Rorschach test should suggest different comparative psychological dimensions at the cross-cultural and intracultural levels. Cultures may not differ the way individuals within cultures differ.

**Bales's Interaction Process Analysis**

Bales’s interaction process analysis (IPA) is a widely used system for coding interpersonal behavior in small groups (see Bales [1968] for a review of the system).

Bales and Hare (1965:240) have published average percentage scores for each of their twelve behavior categories in twenty-one different types of groups (e.g., juries, therapy groups, mediation boards). I used this data to calculate the across unit correlations among the IPA categories, using groups instead of cultures as units.

At the intragroup level there is little published data on the relations among the IPA categories. Fortunately, D’Andrade (1970:26) has reproduced intragroup correlations among seven modified IPA categories which he discovered in an unpublished study by Mann (1959). The correlations were calculated for individuals within one group in each of two different group conditions (socioemotional and task-instrumental).

Mann’s correlations are based on percentage scores for individuals and come from immediate scorings of behavior. He modifies the twelve IPA categories to seven by combining *shows solidarity* and *shows agreement*, *gives opinions* and *gives information*, *asks for*
opinions, asks for information, and asks for suggestions, and shows antagonism and shows disagreement.

I modified the published mean percentage scores for groups (Bales and Hare 1965:240) in the same way, and calculated the cross-group correlation matrix among the seven categories. The resulting matrix was compared to the two intragroup matrices using the cell ranking method discussed above. The two intragroup matrices are not very similar to the cross-group matrix. The correlations for the socioemotional and task-instrumental conditions are .18 and .20 (approximate rₚ) respectively. Again the patterning of within and across unit variation does not coincide.

THE INDEPENDENCE OF BETWEEN AND WITHIN UNIT VARIATIONS

FORMAL CONSIDERATIONS

There are no formal grounds for expecting cross-cultural and intracultural correlations to correspond. A discussion by Hills (1957) makes it apparent that a series of negative intracultural correlations is entirely compatible with a positive cross-cultural correlation using the same variables and the same subjects.

Hills (1957) discusses a procedure for estimating the overall degree of correlation between two variables from information about the degree of correlation between the two variables within several groups of individuals.

The problem is identical to one of estimating the degree of association between two variables across cultures from knowledge of their association within each culture.

In figure 1 (a modification of a diagram presented by Hills 1957: 131) the discrepancy between cross-cultural and intracultural variations is illustrated. A scatterplot of hypothetical individuals from each of four cultures (C₁–₄) on two variables (severity of weaning and oral explanations of illness) is presented.

An inspection of figure 1 indicates that (1) the cross-cultural correlation between severity of weaning and oral explanations of illness is positive (the dotted ellipse), and (2) the intracultural correlations between severity of weaning and oral explanations of illness are all negative (the solid ellipses). Across unit and within unit variations are different for the same variables and same subjects.
Severity
of
Weaning

Oral Explanations of Illness

**Figure 1.** Scatterplot showing a hypothetical relationship between cross-cultural and intracultural correlations among the same variables.

**Between and Within Unit Discrepancies: Interpretations**

Theoretical variables that are cross-culturally valid but lack valid intracultural indicators (see tables 1–3) lend themselves to two interpretations: (1) the theoretical variable is not satisfactory for intracultural comparisons, and (2) a different, as yet undiscovered, but discoverable set of indicators is needed at the intracultural level. The first interpretation implies independence between cross-cultural and intracultural variables. The second interpretation implies independence between cross-cultural and intracultural operational indicators. Logically there may be no way to distinguish the independence of theoretical variables from the independence of assessment procedures. One can not know that an unmeasured variable is unmeasurable, and it seems unlikely that good theoretical variables will be abandoned _simply_ because they have not _yet_ been discovered at some level of analysis.

Certainly there are many measurement properties that could ex-
plain the discrepancy between cross-cultural and intracultural correlations; for example, the lack of control over extraneous variables at a presumably more attenuated intracultural range of variation.

It is also true that different yet equivalent intracultural indicators of valid cross-cultural variables such as nurturance versus egoism may yet be discovered. Still, I think we can entertain the possibility that the search for such indicators will prove futile in some cases. Then the question arises: how can a characteristic that is true of a group or culture in comparison with other groups or cultures not be true of the individuals within those groups or cultures in comparison with one another?

I do not have a theory to help distinguish unsatisfactory variables from inadequate indicators nor even a full list of mechanisms that might account for the intraunit irrelevance of valid cross-unit variables. By listing a few such mechanisms briefly and tentatively, however, some credence can be lent to the possibility that valid cross-unit variables are inappropriate within the units.

1. The cross-unit variable has a locus that is absent within the unit. For example, in the small group data cited above the cross-unit correlation between asking for information, opinions, and suggestions and giving information and opinions is +.37. The average within group correlation between these two variables is −.30. In other words, as the amount of asking for information and opinions increases in a group, the amount of giving of information and opinions increases. Yet within a group those individuals who ask for information and opinions are not those who give information and opinions.

   At the cross-group level the relevant locus of behavior is role-role where the roles are reciprocal and always occupied. Opinion asking elicits opinion giving. At the intragroup (individual) level the relevant locus is the person-role where roles present themselves as exclusive alternatives and the person who is the asker is not the giver.

2. The cross-unit variable is the intraunit background factor that makes it inappropriate within the unit. For example, Whiting and Whiting (1971:38–39) discover a positive cross-cultural relationship between the importance of herding as a task assigned to boys and responsible social behavior in boys. At the intracultural level the relationship disappears. The Whitings suggest that herding cultures place such a high value on responsible social behavior that even nonherding boys in herding cultures learn to be responsible.
Kendall and Lazarsfeld (1955:295) give another example. Across military units in the American armed forces satisfaction with promotion is negatively related to promotional opportunities while within units the relationship is positive. The authors explain that the promotional climate of the unit affects the attitude of its personnel. "The promoted man in an advantageous unit enjoys his promotion less and the nonpromoted man resents his setback more." Hence the negative across unit correlation.

3. The cross-unit variable is dependent upon a cross-unit distribution of its locus. For example, nurturance versus egoism is a cross-cultural role-conditioned variable related to the amount of time spent in baby-sitting tasks. Cultures differ considerably in the frequency with which children typically interact with infants. Intraculturally, however, most members of a society may have similar amounts of access to the relevant role or may have a degree of access that is above some crucial threshold.

The reader should also consult the list presented by Minturn and Lambert (1964:255–257).

**An Implication for Comparative Cultural Research**

I conclude my discussion of within and across unit discrepancies by suggesting that under certain conditions cross-cultural and intracultural research are not mutually relevant. There are two modes of comparative cultural research (see Sears 1961:445–446). In the first mode a cross-cultural population sample is selected and correlations across this sample of cultures are used to test hypotheses and discover universal dimensions of variation.

In the second mode a comparative intracultural replication is performed and the correlations within each of a number of cultures are used to test hypotheses and discover universal dimensions of variation.

Our discussion has raised the possibility that, for reasons other than "problems of measurement," all known cultures and individuals within each of all known cultures may not differ in the same ways. Valid indicators of a theoretical variable may be discovered across a representative sample of cultures without being discoverable within any of them, or may be discovered within each of all cultures without being discoverable across them.
In other words, it is just possible that the psychological variables that will one day describe differences among cultural communities are not the same as those that will describe differences among individuals within any or all cultural communities, and vice versa.

REFERENCES


