

# The Soft Side of Hard Data in the Study of Cultural Values

RICHARD A. SHWEDER

## Introduction: A Blast from the Past

I originally wrote this essay in September 1970 while I was engaged in long-term fieldwork in the Hindu Temple Town of Bhubaneswar in Orissa, India. I prepared the document (and then just filed it away until now) as a finger exercise of sorts concerned with the soft side of hard data and the subjective or discretionary aspects of quantitative data analysis and interpretation. But looking back I can see that this previously unpublished demonstration piece marked the starting point of my disenchantment with abstract value questionnaires as an objective method for studying cultural differences in folk understandings of the social order as a moral order. It thus seems appropriate and even timely to revive and revivify the essay now, because the field of moral anthropology has returned to center stage in cultural anthropology in recent decades (Fassin 2012).

That phrase or banner—moral anthropology—is, of course, semantically ambiguous and points in two directions. On the one hand it points in the direction of a moralizing (or even moralistic) anthropology that is critical, judgmental, and normative and concerned with studies

of, e.g., inequality, poverty, violence, social justice, and human rights. On the other hand, the phrase points in the direction of the anthropology of morality dedicated to positive science, comparative studies of values, moral reasoning, moral sentiments, and cultural narratives.

Not surprisingly, the ancestral heroes of contemporary moral anthropologists are themselves ideologically diverse: Emile Durkheim, Max Weber, Bronislaw Malinowski, Kenneth Read, Karl Marx, E. P. Thompson, Michel Foucault. Notably, mixed methods (or perhaps more accurately, mixed-up methods) are the research methods employed by card-carrying moral anthropologists. A list of those methods would include archival analysis of legal cases and dispute resolutions, behavioral documentation of folkways and customary practices, structured interviewing, value questionnaires, conversational analysis, cognitive experiments, media analysis, recording neuronal activity in the human brain, etc.

In any case, back in 1970 when I wrote “The Soft Side of Hard Data,” I was investigating concepts of the person in Orissa, India. That research later morphed into a comparative project on folk conceptions of the social order as a moral order, which led to the empirical and theoretical development of a conception of the moral domain known as “the Big Three.” This comparative approach to the study of morality argues that on a global scale, within any culturally diverse society, and across institutional domains within any cultural group, moral judgments about the right thing to do under such-and-such circumstances are linked to a variety of diverse terminal goods or ultimate values, which come in three broad kinds: moral judgments ultimately grounded in an ethics of autonomy (with its focus on values or moral goods such as harm, rights, and justice and an emphasis on the self as a preference structure with wants entitled to satisfaction); moral judgments ultimately grounded in an ethics of community (with its focus on values or goods such as duty, loyalty, interdependency, and respect for hierarchical authority and an emphasis on the self as a responsible and self-controlling status bearer with a social role to play within a community); and moral judgments ultimately grounded in an ethics of divinity (with its focus on purity, sanctity, cleanliness, and the connection between the natural order and the sacred order and an emphasis on the self as a potentially elevated or dignified token of some divine, transcendent, or higher order of things).

Given that interest in the study of morality and conceptions of the person, I was busy in 1970 reading about quantitative methods for the study of values, such as the ranking and rating procedures associ-

ated with abstract value questionnaires (Hays 1963). Value questionnaires ask informants from different cultural groups, ethnic traditions, or social categories to make ratings or rankings of abstract words or context-free stand-alone value concepts, for example, “Which is more important to you, beauty or wealth, freedom or equality, security or privacy, honesty or cleanliness?” Or, alternatively, how important is “obedience” on a five-point scale? How about “loyalty”? How about “self-control”? How about “justice”?

Looking back now, I wish Roy D’Andrade’s little-read and astonishingly underappreciated 2008 volume *A Study of Personal and Cultural Values* had been available to me when I was in graduate school. His brilliant book was written much later, when social psychologists were drawing attention to modes of East Asian thought and contrasting them with Anglo-American modes of thought. In the thick of that widespread contemporary discussion of East versus West differences in cognition, D’Andrade developed a quantitative study of what he called the personal values of informants from three groups: American undergraduates at the University of California, Vietnamese refugees to the USA, and Japanese respondents living in Japan. Writing against the current of a cultural psychology focused on East versus West differences, his main take-home message was that he had a very hard time finding significant cultural or group-based differences in the expressed personal values of his American, Vietnamese, and Japanese informants. Between-group variance in the endorsement of abstract value words and phrases was minimal. For eager readers of the anthropology of morality, there was nothing much to write home about concerning cultural differences.

Fortunately Roy D’Andrade is a brilliant culture theorist as well as a virtuoso quantitative methodologist. So, confronted with his findings he went on to draw a useful theoretical distinction, which I would commend to researchers in cultural psychology and moral anthropology. He distinguished the study of personal values (tell me which is more important to you, cleanliness or honesty, liberty or justice?) from a different unit of analysis focused on values (and inevitably beliefs) with respect to something in particular (such as—to pick a not-so-random example—being the chairman of the Center for Psychosocial Studies at UCLA or the Editor of *Ethos*), which he dubbed institutional values. To get a better sense of this distinction: the study of “institutional values” is not an investigation of an individual’s evaluation of the relative importance of abstract goods (such as liberty, justice, loyalty, and personal sanctity). Instead it is the study of all the ideas about what is true,

good, and instrumental, revealed and made manifest through speech when an informant is questioned about the obligations and expectations associated with role or status-based participation in the customary practices and local institutions which are familiar to the informant from living in his or her own particular society—for example, all the beliefs and values associated with being a widow in a Hindu temple town in India. (I will have more to say about cultural differences in widowhood in a moment.)

Looking back to 1970, I think the unit of cultural and mental analysis described as institutional values by Roy D'Andrade was probably available (at least in theory) to students of John Whiting through a cognate concept that Whiting and Irvin Child had discussed in their 1953 collaboration *Child Training and Personality*. There Whiting and Child managed to combine the comparative study of routine cultural practices (or folkways) and the comparative study of individual mentalities into a single unit of analysis called the custom complex. Simply put, Whiting and Child state that the custom complex “consists of a customary practice and of the beliefs, values, sanctions, rules, motives, and satisfactions associated with it.” In my view the study of institutional values is another way of describing the study of the custom complex.

Although Whiting and Child introduced the idea of the custom complex in 1953, its theoretical implications were not widely or fully appreciated at the time. For the most part the idea was not taken up or carried forward by psychological anthropologists working in the 1950s. Nor did the custom complex become a unit of analysis for social psychologists working on the development of value questionnaires. Indeed, it was not until the 1980s and 1990s, when Weisner and others developed a research agenda for the study of the ecological niche (Weisner 1984, 1996; also 2001, 2002) and a so-called practice approach or activity-setting approach took hold among some child development researchers (Goodnow, Miller, and Kessel 1995), that Whiting and Child's conception gained some limited currency and appeal. The study of the custom complex also bears some resemblance to the founder of American social psychology Kurt Lewin's idea of a personal life space (Lewin 1943) and fully anticipates the French sociologist Bourdieu's now-popular idea of a habitus (Bourdieu 1972, 1990). It is the study of the mental side of social habits, which are always bound to particular institutions and social roles in specific cultural settings. In my view, paying more attention to this broader unit of analysis is

a good way to make progress in the study of moral anthropology and cultural psychology.

### Mixed-Up Methods: Exploring Cultural Differences in Values

The recognition of the bounded or qualified nature of social science generalizations, including the extent to which one's picture of social and psychological reality is dependent on the particular methods one uses to generate "data" so as to interpret what is real (so-called method variance—Campbell and Fiske 1959), goes hand in hand with the study of the custom complex, the ecological niche, and institutional values. That recognition—that social science generalizations are typically narrow in scope—is not new. The philosopher of science Ernest Nagel made the observation in his 1961 book *The Structure of Science* and he cites a reference from 1934. He writes (459): "The conclusions reached by the controlled study of sample data drawn from one society are not likely to be valid for a sample obtained from another society. Unlike the laws of physics and chemistry, generalizations in the social sciences therefore have at best only a severely limited scope, limited to social phenomena occurring during a relatively brief historical epoch within special institutional settings." He makes the point that human actions are mediated by local "technologies and traditions," and although Nagel himself leaves open the possibility (and hopes) that generalizations of broad scope will be discovered, he avers that "the possibility must certainly be admitted that nontrivial but reliably established laws about social phenomena will always have only a narrowly restricted generality" (460). Starting one's research on values by focusing one's interview questions on bounded units such as a particular custom complex, or a particular ecological niche, or the particular behavior in the context of a particular social institution is one way to acknowledge that organizational feature of human social life.

In his 2008 book Roy D'Andrade milked the study of informant judgments about abstract value words and phrases for all they are worth, and came up with very little by way of cultural differences. His research program is especially instructive for those of us interested in psychological differences across human populations because he went into the research fully aware of the many problems with value questionnaires yet not yet prepared to abandon them as a method for the study of values. Thus he writes:

There are well known problems with questionnaires. The same words mean different things to different people. Translations are imperfect. People, however honest their report, do not always respond to the words for things the way they respond to things themselves. Someone may think they value something highly when presented with words—for example, how much do you value peace and quiet?—yet when presented with lots of peace and quiet may find they do not value it as highly as they thought. And people may simply not be able to answer some questions—they just do not know how much they value X and may never know. Or they may be profoundly ambivalent about something, and both value it and disvalue it, so that no single rating covers the situation. Despite all these problems, with respect to efficiency and efficacy, there is much to be said in favor of questionnaires for the study of values. Observation of the choices someone makes cannot tell us what that person thinks or what he or she feels is good. The most efficient way to find out what people think is to ask them. One can observe people smoking cigarettes but they may or may not think smoking is a good thing. (D'Andrade 2008, 13)

(Parenthetically, one does wonder whether declaring one's general commitment to an abstract value—such as equality or piety or social justice—and having one's identity as a person associated with such a public commitment is really a familiar or basic feature of cultural, political, and psychological life in all societies. The method itself and its very mode of questioning may presuppose a bit too much about fundamental psychological and linguistic practices.)

D'Andrade's volume is full of such pithy methodological observations about value questionnaires and theoretical insights relevant to the study of similarities and differences in psychological functioning across cultural groups of the sort pioneered by generations of psychological anthropologists. Nevertheless, his main take-home message and empirical finding after years of carefully conducted quantitative research is that Japanese, Vietnamese, and Anglo-Americans do not differ very much in their values, at least not when they are systematically asked to judge abstract value words or phrases.

I would describe his finding as a potential point in favor of the principle of "method variance," discussed by Campbell and Fiske (1959). The data one generates on similarities or differences in values across cultural groups is not independent of the procedures one used to generate the data. The study of informant ratings or rankings of abstract and decontextualized value words or phrases may not be the best way to get at cultural differences in the beliefs and values that give meaning and value to the moral worlds institutionalized in different cultural traditions.

Implicit in the ideas of a custom complex, an ecological niche or an institutional value is the imperative to begin one's research on other societies with fieldwork that documents the local institutional world and customary or habitual practices of members of a distinct cultural group. It is an invitation to then structure one's interviews in such a way that they provoke a discussion of the values and beliefs associated with local institutions or ecological niches, using the native's language to get at the beliefs and values that are important to being, for example, a mother, a teacher, the conductor of an orchestra, a widow; in other words, focus the interview on the folk custom or institution in which a person in such-and-such status plays a significant part rather than treating the informant as a stand-alone individual with a general motivating ideological commitment to abstract values regardless of context. Many Americans on the liberal left who say they value freedom do so because they have in mind concrete, fact-based institutional contexts, for example, the freedom of the *New York Times* to publish classified government documents such as the Pentagon Papers, which activates some master metaphor such as the free market of ideas. But if you press on and concretize the abstract value in a different way, you may discover they are not so sure they value freedom when it is comes down to a free market of goods and services and the institutional context is the workplace, the freedom of employers and employees to enter into labor contracts and the right of skilled workers to be hired for a job regardless of union membership.

Context matters and it needs to be built into our value interviews. There are many ways to do this. After thinking about my own work on cultural differences in moral judgments in the Temple Town of Bhubaneswar, Orissa, India, and Hyde Park, Illinois, USA, in the light of Weisner's notion of an ecological niche and D'Andrade's notion of institutional values, it seemed fortunate to me that the interviewing we did in the early 1980s was role-based and embedded in local customary practices.

Here is an example drawn from that work (Shweder, Mahapatra, and Miller 1990; also see Shweder et al. 1997). Below are two interviews contrasting institutional values in the Hindu Temple Town of Bhubaneswar, Orissa, India, and in the secular community of Hyde Park, Illinois, USA. The custom complex in this instance is widowhood. The interview examines the respondents' ideas about what is true, good, and instrumental with respect to marriage and widowhood. Those ideas are elicited by structuring a series of interrogatives about a particular pattern of behavior by a hypothetical widow. Consider the fol-

lowing behavior: “A widow in your community eats fish two or three times a week.”

First, the interview in the Hindu Temple Town:

Is the widow’s behavior wrong? (Yes. Widows should not eat fish, meat, onions, or garlic, or any “hot” foods. They must restrict their diet to cool foods, rice, dhal, ghee, vegetables).

How serious is the violation? (A very serious violation. She will suffer greatly if she eats fish.)

Is it a sin? (Yes. It is a great sin.)

What if no one knew this had been done? It was done in private or secretly. Would it be wrong then? (What difference does it make if it is done while alone? It is wrong. A widow should spend her time seeking salvation—seeking to be reunited with the soul of her husband. Hot foods will distract her. They will stimulate her sexual appetite. She will lose her sanctity. She will want sex and behave like a whore.)

Would it be best if everyone followed the rule that widows should not eat fish? (That would be best. A widow’s devotion is toward her deceased husband—who should be treated like a god. She will offend his spirit if she eats fish.)

In the United States widows eat fish all the time. Would the United States be a better place if widows stopped eating fish? (Definitely it would be a better place. Perhaps American widows would stop having sex and marrying other men.)

What if most people in India wanted to change the rule so that it would be considered all right for widows to eat fish? Would it be okay to change the rule? (No, it is wrong for a widow to eat fish. Hindu dharma—truth—forbids it.)

Do you think the widow who eats fish should be stopped from doing that or punished in some way? (She should be stopped. But the sin will live with her and she will suffer for it.)

Next consider the parallel interview conducted in Hyde Park, Illinois, with a secular informant:

Is the widow’s behavior wrong? (No, she can eat fish if she wants to.)

How serious is the violation? (It is not a violation.)

Is it a sin? (No!)

What if no one knew this had been done? It was done in private or secretly. Would it be wrong then? (It is not wrong in private or public.)

Would it be best if everyone followed the rule that it is all right for a widow to eat fish if she wants to? (Yes, people should be free to eat fish if they want to. Everyone has that right.)

In India it is considered wrong for a widow to eat fish. Would India be a better place



if it was considered all right for a widow to eat fish if she wants to? (Yes, that may be their custom but she should be free to decide if she wants to follow it. Why shouldn't she eat fish if she wants to?)

What if most people in the United States wanted to change the rule so that it would be considered wrong for a widow to eat fish? Would it be okay to change it? (No, you can't order people not to eat fish. They have a right to eat it if they want to.)

Do you think the widow who eats fish should be stopped from doing that or punished in some way? (No!)

I submit that one learns far more about cultural differences in moral understanding and folk conceptions of the social order as a moral order from interviews of this type than by asking devout Hindus in India and devout secularists in the USA whether and to what extent loyalty or freedom of choice is good. I hope I have said enough by way of introducing my flash from the past. Below is that 1970 demonstration study expressing doubts about the objectivity and usefulness of value questionnaires as a method for understanding cultural differences in the moral domain.

### The Soft Side of Hard Data (1970)

This is a note about uncertainty in the analysis and interpretation of some social science data. I shall show how "hard data" on values in American society may be statistically analyzed so that for every one of four plausible and mutually exclusive interpretations of the data, three others may be advanced that are also plausible and compatible with the facts.

There are three reasons for considering the analysis and interpretation of data indeterminable.

1. There is indeterminateness between the data and the analysis. Each interpretation is inferred from a statistical analysis that draws or ignores some distinctions that potentially exist in the data. The data do not provide us with objective criteria for deciding which distinctions to heed and which to disclaim in the data when performing the statistical analysis. Thus what appears to be a conclusion drawn from the facts is rather an interpretation only loosely constrained by the data.
2. There is indeterminateness between the results of the statistical analysis and the interpretation. Each interpretation is compatible with the results of a statistical

analysis that may be performed on the data. None of the interpretations necessarily follow from the analysis.

There is indeterminateness in evaluating the best interpretation. Any one of the interpretations may be correct and there is no way to judge from the data per se.

## Two Faces of Values Data

It is possible for members of diverse social categories to disagree about the relative importance of every value in their value system, and yet almost perfectly agree on the hierarchical ordering of their value system as a whole.

So for any set of value statements, disagreement over the relative importance of each value considered alone may be statistically significant while, simultaneously, agreement about relative importance may be statistically significant for the system as a whole. In fact as the number of values about whose relative importance people can disagree increases, the extent of agreement about the system as a whole may approach perfection.

This paradox makes it possible to analyze values from two points of view that lead to strikingly different interpretations of the degree of similarity of the value systems of members of various categories in society.

### *A First Analysis and Interpretation*

In a study entitled "Values as Social Indicators of Poverty and Race Relations in America," Rokeach and Parker (1970, 98) asked a large sample of Americans to rank two sets of eighteen value statements in terms of their importance as guiding principles in the informant's daily life. The informants varied considerably in the level of their income and education, and in their race.

The results of the study were presented in eight tables. The first and second tables gave the median and composite rank of each of the values as judged by members of different income categories ranging from under \$2,000 a year to \$15,000 and over. The third and fourth tables gave the same type of information for values as ranked by members of different educational categories ranging from zero to four years of

schooling to a graduate-school level of education. The fifth and sixth tables gave the ranks of the values as judged by blacks and whites randomly sampled. The seventh and eighth tables gave the ranks of the values for blacks and whites matched for income and education.

Rokeach and Parker asked the following question of their data: how many of these eighteen values in each of these tables are significantly different in the relative rank assigned to them by members of different social categories? They answered the question by applying the median test to each value in the tables. In table 6.1, I have reproduced some of their data for income categories. The table shows the median and composite rankings of eighteen values by informants of diverse socio-economic status from poverty to affluence. The results of applying the median test indicate that eleven of the eighteen values are significantly different among income groups.

The results of the statistical analysis performed by Rokeach and Parker are as follows: (1) twenty of the thirty-six values in the two income tables show significant differences among income categories, (2) twenty-five of the thirty-six values in the two education tables show significant differences among educational groups, (3) fifteen of the thirty-six values in two of the racial tables show significant differences between blacks and whites, but this number is reduced to seven when the informants are matched for income and wealth.

Based on the results of their statistical analysis, Rokeach and Parker make an interpretation about the extent of difference in the patterning of values by informants having various social characteristics in American society. Their interpretation is an evaluation of the number of significant differences out of thirty-six possibilities that appear in each pair of income, educational, and racial tables (106, 108). Rokeach and Parker construct the following interpretation: (1) there are significant and pervasive variations in value systems, associated with differences in socioeconomic status in America (106), and the values of the poor differ significantly from the values of the more affluent segments of our society (110). (2) Most of the differences in values between blacks and whites in America can be attributed to differences in socioeconomic status and not to distinctive racial cultures (108).

Their interpretation indicates a tendency toward a “culture of poverty” in America and suggests that black culture is different from white culture only to the extent that a greater proportion of blacks are poor, and not because of a different cultural heritage and tradition. I would add another interpretation to their statistical analysis, which I think

Table 6.1. Value medians and composite rank-orders for informants varying in income (N=1,325)\*

Values	Under \$2,000		\$2,000-\$3,999		\$4,000-\$5,999		\$6,000-\$7,999		\$8,000-\$9,999		\$10,000-\$14,999		\$15,000 & Over		Median $\chi^2$	Test $p =$
	N=139		N=239		N=217		N=249		N=178		N=208		N=95			
	MD.	RNK.	MD.	RNK.	MD.	RNK.	MD.	RNK.	MD.	RNK.	MD.	RNK.	MD.	RNK.		
Ambitious	8	6	6.9	3	6.1	2	6.8	3	6.6	3	5.8	2	6.4	3	8.33	0.215
Broad-minded	8.6	8	7.2	4	8.1	8	8.1	6	7.1	4	6.4	4	7	4	11.92	0.064
Capable	9.5	10	10.5	14	9.8	11	9.3	11	9.8	11	8.4	7	8.8	8	14.98	0.02
Cheerful	9	9	8.6	9	10.6	14	10.3	12	10.7	12	10.2	11	11.3	14	16.33	0.012
Clean	6.4	2	7.3	5	8	7	8.6	8	9.3	10	10.4	12	14.4	17	72.35	0.001
Courageous	7.5	5	8.1	8	8	5	7.5	5	7.4	6	8	5	7.2	5	2.37	0.825
Forgiving	6.4	3	6.5	2	7.3	4	6.3	4	7.4	5	8.1	6	10.7	12	20.23	0.003
Helpful	7.1	4	7.4	6	8	6	8.2	7	8.9	8	9.3	8	9.1	9	17.69	0.007
Honest	3.3	1	3.7	1	3.4	1	3	1	3	1	3.4	1	3	1	4.71	0.582
Imaginative	15.2	18	15.8	18	15.6	18	15.9	18	15	18	14.6	18	11.4	15	29.32	0.001
Independent	10.5	14	10.3	12	10	12	10.7	14	11.4	13	11	13	8.3	6	5.54	0.476
Intellectual	13.9	16	13.4	16	13.3	16	13.6	16	13.1	15	12.1	15	8.6	7	25	0.001
Logical	15.2	17	14.8	17	14.7	17	14.1	17	14	16	12.8	16	10.9	13	36.56	0.001
Loving	10	11	10.3	13	9.5	10	9.1	10	8.8	7	10.2	10	9.8	10	2.47	0.872
Obedient	12	15	12.4	15	13.3	15	13.2	15	14.2	17	14.3	17	15.3	18	27.06	0.001
Polite	10.4	13	10.2	11	10.2	13	10.4	13	11.4	14	11.2	14	13.2	16	26.11	0.001
Responsible	8.2	7	7.8	7	7.1	3	5.8	2	6	2	6	3	5.9	2	33.51	0.001
Self-controlled	10.2	12	9.9	10	9.2	9	9	9	9	9	9.7	9	9.9	11	3.21	0.782

\*These data come from Rokeach and Parker (1970, 102) and is table 2 in their study.

the authors would endorse. Since Rokeach and Parker assume that a person's values have social consequences (98) (I take this to mean effects on behavior), their analysis might suggest that (3) behavioral differences between the rich and poor and educated and uneducated can be related to differences in their value preferences.

### *A Second Analysis and Interpretation*

I have reanalyzed Rokeach and Parker's data utilizing different statistics. I have constructed an interpretation of the data that is completely different. I asked the following question: To what degree do members of these various social categories agree in the way they hierarchically rank all the eighteen values in terms of the variable "important to me as a guiding principle in my daily life"? To answer the question I applied Spearman's Rank Order Correlation Coefficient ( $r_s$ ) and Kendall's coefficient of concordance ( $W$ ) to the eight tables.

The results of the analysis are as follows:

1. When Kendall's coefficient of concordance is applied to the composite ranks of values across all income categories in my table 1,  $W =$  approximately  $+0.84$ , and the average correlation is high. Average  $r_s =$  approximately  $+0.81$ . The hypothesis that income groups do not agree on the patterning of values must be rejected ( $\chi^2_{17} = 99.96$ , significant beyond  $.001$ ).
2. When Kendall's coefficient of concordance is applied to the composite ranks of values across all educational categories in Rokeach and Parker's table 4,  $W =$  approximately  $+0.73$ , and the average correlation is high. Average  $r_s =$  approximately  $+0.68$ . The hypothesis that educational categories do not agree on the patterning of values must be rejected ( $\chi^2_{17} = 86.87$ , significant beyond  $.001$ ).
3. In Rokeach and Parker's table 1 on income differences, the composite rank of values for the most extremely different income categories, namely under \$2,000 and \$15,000 and over, are highly correlated,  $r_s =$  approximately  $+0.65$ .
4. In Rokeach and Parker's table 3 on educational differences, the composite ranks of values for informants with extremely different educational backgrounds, namely college graduates and those with only zero–four years of education, are highly correlated,  $r_s =$  approximately  $+0.63$ .
5. The composite ranks of values for blacks and whites are highly correlated in Rokeach and Parker's tables 5, 6, 7, 8. Controlling for income and education *does not* increase the correlation of the system of ranks between blacks and whites. The correlations in the four tables are  $r_s =$  approximately  $+0.81$ ,  $+0.88$ ,  $+0.84$ ,  $+0.86$ , respectively. The highest correlation is for their table 6, before the informants were matched for income and education.

The results of this analysis lead me to construct a very different interpretation. I conclude that: (1) patterning of values by Americans is relatively invariant across differences in wealth, education, and race. (2) Any differences in behavior that may exist among members of different races, the rich and poor, the educated and uneducated, have little to do with the values they profess. (3) The value system remains stable while behavior varies with one's position in the social structure. Therefore much of behavior is controlled by one's situation and not one's beliefs about desirable behavior or future goals. In America the value system does not change very much as you increase your income, attain educational degrees, or change your color, or as you exhibit the behavior appropriate to your new status.

### *An Interlude in Middle Earth*

The same quantitative data lent themselves to two entirely divergent interpretations as a result of the statistics used in the analysis. By applying the median test and Spearman's rank order correlation coefficient to some hypothetical data, it is easy to understand how such opposite results can be produced.

Assume some anthropologist wandering through Middle Earth were to discover a society consisting of two hundred members divided equally between two social groups, the dwarves and the elves. After much painstaking effort the anthropologist was able to formulate all the relevant values of the culture, and these were ten in number. Every member of the society expressed these and only these values. But the dwarves and the elves disagreed on how important each value was, relative to all the others.

The anthropologist then asked every member of the society to rank the ten values and found that all the dwarves gave the same rank to each value, and all the elves gave the same rank to each value, but the two groups never agreed on the precise rank of each value. The composite ranks for the two groups are shown in table 6.2.

If these hypothetical data are analyzed with the median test, such as that applied by Rokeach and Parker, we perform a separate analysis of each of the ten values. For each value we compare the distribution of all the dwarves' ranks and all the elves' ranks with the grand median rank.

With such a procedure we can test the hypothesis that the distribution of ranks around the grand median is significantly different for the two groups with respect to value 1. The grand median for the two

Table 6.2. The ranking of ten values by dwarves and elves: hypothetical data

	Value statement number									
	1	2	3	4	5	6	7	8	9	10
Dwarves composite rank (n=100)	1	2	3	4	5	6	7	8	9	10
Elves composite rank (n=100)	2	1	4	3	6	5	8	7	10	9

Table 6.3. The median test on any value in Table 6.2

	Dwarves	Elves
Above median	100	0
Below median	0	100

groups is 1.5. If we then sort the informants’ ranks in relation to the grand median we get the two-by-two table in table 6.3.

The results are obviously significant. There is a significant difference between dwarves and elves in their precise ranking of value 1. The same results are produced by applying the median test to the other nine values. The results of a median test analysis indicate that dwarves and elves differ significantly on every single value in their Middle Earth culture.

If we apply the Spearman rank order test to the same hypothetical data, we perform one analysis on the whole system of ranks of the two groups. We only wish to know the extent of agreement between the two hierarchies of ranks. We take into account the number of values being ranked and the amount of difference in the rank assigned to each value by the two groups. We assess the degree of agreement, nonagreement, or disagreement in relation to the mathematical possibilities for maximum agreement or disagreement. If we apply the  $r_s$  to the data in table 6.2, the value systems of elves and dwarves correlate almost perfectly,  $r_s =$  approximately  $+ .94$ . And as the number of values about which elves and dwarves can disagree on relative importance increases, the correlation of the value systems approaches greater perfection. With eighteen values arranged according to the system in table 6.2, the  $r_s =$  approximately  $+ .98$ . The correlation continues to approach unity as the number of values disagreed on enlarges!

### *A Third Analysis and Interpretation*

With this insight from Middle Earth, let us return to Rokeach and Parker’s data on American values. It is mathematically possible for very

Copyright © 2016. University of Chicago Press. All rights reserved.

high correlations between rankings to coexist with significant disagreement on the rank of every single value. This fact allows for the construction of a third interpretation of the data. The same data can be analyzed so that: (1) among members of various income and educational categories, forty-five significant differences over the placement of individual values are found out of a possibility of seventy-two differences, (2) the rankings of values of all informants are highly correlated.

The interpretation that is constructed from these results of combining the median test, and the Spearman  $r_s$  and Kendall W analyses is as follows: (1) There is truly an American value system, one that all segments of the population share. (2) There are also a large number of significant points of disagreement over the placement of individual values within the system. (3) The sharing of the value system is considerable and indicates that, in spite of differences in social status, we are all members of one somewhat integrated society. (4) The differences in values are considerable and are related to the variations of social status of members of our society.

This interpretation is appealing because it conforms to our common sense notion that two objects or systems can be both similar and different at the same time (for example, the colors blue and green), and because it seems to account for the results of both statistical analyses. The large number of differences over the placement of values discovered by the median test is combined with the high correlations between the total system of ranks discovered in the concordance analysis. The interpretation is compelling because it encompasses both types of statistical results.

#### *A Fourth Interpretation*

The results of a further statistical analysis will provide a basis for a fourth interpretation. The number of significant differences in values (as judged by the median test), and the size of the correlation of any two or more systems of ranks (as judged by  $r_s$ ), are independent measures. Nearly identical correlations in the overall ranking can exist between rankings where every value is significantly different, or alternatively where none of the values are significantly different. An identical number of significant differences between individual values can exist between rankings that have a perfect negative correlation or, alternatively, a nearly perfect positive correlation.

In table 6.4, assume that each ranking is the ranking of members of different social categories. Assume that agreement within each social



Table 6.4a and Table 6.4b. A comparison of four rankings of values showing the independence of results of the median test from results of Spearman’s correlation coefficient

	Values										
Ranking 1	1	2	3	4	5	6	7	8	....	17	18
Ranking 2	1	2	3	4	5	6	7	8	....	17	18
Ranking 3	2	1	4	3	6	5	8	7	....	18	17
Ranking 4	18	17	16	15	14	13	12	11	....	2	1

  

	Median test number of significant differences in individual rank	rs correlation of two rankings
Rankings 1 & 2	0	+1.00
Rankings 1 & 3	18	+0.98
Rankings 1 & 4	18	-1.00

category about that ranking is complete. Rankings 1 and 2 and rankings 1 and 3 have nearly identical correlations (+1.00 and +.98) and the extreme opposite number of differences in individual ranks (0 and 18). Rankings 1 and 3 and rankings 1 and 4 have the same number of individual differences in rank (18) and completely opposite correlations (+.98 and -1.00).

The same number of differences in individual ranks can exist in systems that are completely unlike as well as in systems that are identical. This means that the median test fails to make an important distinction among those values that are found to be significantly different in their individual rank. It does not distinguish among values in terms of the extent of difference between their ranks. It is precisely the extent of difference that determines the extent of correlation of the two rankings.

The Spearman  $r_s$  statistic is defined as  $1 - (6(\sum_1 D_1^2) \div N(N^2 - 1))$ . The two variables in the formula are  $D$ , the size of the difference in rank of each object in the two rankings, and  $N$ , the number of objects ranked. In rankings of the same number of objects,  $N$ , of course, is a constant. The only variable is  $D$ , the size of the difference in ranks for each object in the rankings.

I have already shown in table 6.4 that the same number of significant differences in the median test can tolerate rankings whose correlations range from -1.00 to +.98. It follows that when analyzing the data on values it is not useful to count the number of significant differences in the median test results. Rather, we should distinguish each individual difference as a contributor to, or a detractor from, the overall correlation of the rankings.

Copyright © 2016. University of Chicago Press. All rights reserved.

If one divides the values that Rokeach and Parker discovered to be significantly different in rank into those that contribute to high correlations in the two rankings and those that reduce the similarity of the two rankings, one gets the following results.

1. Forty-one out of sixty-seven cases of significantly different values (60 percent of their findings) have an average difference in rank of three ranks or less. Rankings of eighteen values between which no value is more than three ranks different have minimum correlations of  $+0.84$  ( $r_s$ ).
2. At least fifty of sixty-seven cases of significantly different values (nearly 75 percent of the findings) have an average difference in rank of four or less. Rankings of eighteen values between which no value is more than four ranks different have minimum correlation of  $+0.71$  ( $r_s$ ).
3. At best only seventeen cases out of a possible 144 are different enough to reduce the size of the correlations found between the rankings of all informants.

On the basis of this analysis one might construct the following interpretation: (1) there is truly an American value system, one that all segments of the population share. (2) There are also a small number of significant points of disagreement over the placement of individual values within the system. (3) The sharing of the value system is considerable and indicates that in spite of differences in social status, we are all members of one somewhat integrated society. (4) The differences in values are negligible but are related to variations in the social status of members of our society.

### *The Four Interpretations*

The four interpretations and their statistical foundations are summarized in table 6.5. The four interpretations of the data are alternative and mutually exclusive interpretations. Each is confirmed by a different set of statistical results with which it can be viewed as compatible. If we accept the reference and sense of any one of them, we cannot accept the reference and sense of any other, nor can we limit ourselves to the set of statistical results with which they are compatible.

I include the "sense" of the interpretation in this reasoning about their mutual exclusiveness because in each interpretation's reference, interpretations one and three are not mutually exclusive. It is possible to have pervasive differences in the value system (interpretation one) and at the same time have sharing. This is the reference of the third interpretation. But in their sense, the two interpretations are quite in-

Table 6.5. Summary of four statistical analyses of the data and their interpretations

Interpretation	Statistical Results
1. Pervasive differences in values associated with socioeconomic differences. Behavior variations related to differences in value preferences.	The median test indicates out of a possible 72 differences, 45 significant differences in the rank of values by informants varying in socioeconomic status.
2. Value preferences invariant across differences in socioeconomic status. Behavioral differences not related to differences in values.	The Spearman $r_s$ and Kendall W indicate high correlations between rankings of all informants regardless of socioeconomic status.
3. A value system shared by all socioeconomic categories but a large number of differences in the placement of individual values. Sharing considerable. Differences considerable.	The two sets of results listed above.
4. A value system shared by all socioeconomic categories but a small number of differences in the placement of individual values. Sharing considerable. Differences negligible.	The high correlations in the Spearman and Kendall tests—and the fact that only 25% of the differences found by the median test—detract from the high correlation of the informants' rankings.

compatible. Considered alone, the first interpretation lays emphasis on, or implies the fact that the rich and poor and the educated and uneducated have different value systems. This implication is completely lacking in the third interpretation.

### *From Data to Analysis*

There is a degree of uncertainty between the data and the statistical analysis. The data do not provide us with a criterion that might aid us in deciding which statistical analysis is interpretatively relevant. The data can always be analyzed in different ways, and more and more statistical distinctions can be added. We might, for example, decide that the only important differences are between values whose difference in rank is large enough to produce negative correlations between two rankings. This would reduce the significant differences to less than five percent of all the values ranked.

But the truth of the interpretation does not increase as the number of statistical distinctions upon which it is based is increased. Some differences in the data may be interpretively irrelevant, as I assumed in the fourth analysis and interpretation. The data themselves remain silent on this issue.

The four interpretations are mutually exclusive alternatives. They are all compatible with aspects of the data as shown in a statistical

analysis. There is no criterion given by the data per se for preferring any one statistical analysis. It follows that the data do not impel us to adopt one interpretation or another. The facts do not speak for themselves. The only criterion that does seem available for selecting one statistical analysis over another is the interpretation that one personally favors, or hopes to discover in the data, or believes deserves to be true.

### *From Analysis to Interpretation*

Uncertainty is not limited to the step from the data to the selection of a statistical technique for its analysis. There is indeterminateness between the results of the statistical analysis and the interpretation. Although each interpretation is compatible with the results of some statistical analysis, not one of the interpretations necessarily follows from those results. I will consider each interpretation in turn.

The first case begins with an analysis that indicates a large number of statistically significant differences in individual rank. It concludes with the interpretation that pervasive differences in the value system are associated with socioeconomic status. I have demonstrated that the number of individual differences in rank that are significantly different can remain fixed while the degree of correlation of the value system can vary from complete disagreement to near complete unanimity. It can be argued that we cannot ignore the extent of correlation of the rankings when we interpret the value rankings of two groups to display pervasive differences. Thus an assessment of the extent of difference cannot be directly inferred from such statistical results.

The second case begins with an analysis that indicates high correlations between the value systems of all informants. It concludes with an interpretation that value systems are invariant across socioeconomic categories. It is the nature of the correlation coefficient that two systems of ranks can be highly correlated and still display statistically significant differences in every value. Who can say from the data themselves whether a statistical difference as small as one or two ranks is conceptually or behaviorally important or not? It is also possible for two rankings to be highly correlated and still have major differences in rank between a small number of values. For example with eighteen values, if rankers agree on the rank of sixteen of the values, the remaining two values can be twelve ranks apart and the  $r_s =$  approximately  $+0.71$ . These two values with large discrepancies in rank may be the key to all the behavioral differences between the social categories. Or they may

be completely irrelevant. The data remain silent on the issue. Thus an assessment of the extent of similarity cannot be directly inferred from these statistical results either.

The third and fourth interpretations do not necessarily follow from the statistical analyses with which they are compatible. The third case begins with a statistical analysis that indicates a large number of statistically significant differences in the placement of individual values between members of various socioeconomic categories, and high correlations between the total rankings of these informants. It concludes with an interpretation that there is considerable sharing of values and considerable differences at the same time. Yet from a strictly statistical point of view the interpretation does not directly follow. Since the number of individual differences in value placement varies independently of the correlation coefficient, it can be argued as above that only a small number of differences are statistically significant. In this case statistical significance will be determined by how much the difference in the individual values detracts from the overall correlation of the rankings of all values. The interpretation that there are considerable differences would then be eliminated.

From the cultural point of view, the third interpretation is only one possible conclusion compatible with its statistical results. The question of whether you place “equality” above “personal pleasure,” or vice versa, may place you into one of two somewhat hostile political camps in our contemporary society. How much you place it above or below in importance really may not matter. But the question of whether you feel that “a world of beauty” is more important than “inner harmony” may have no implications conceptually, politically, or socially to members of our culture. That is to say, we cannot assume that all the values ranked are weighted equally in their social implications in daily life: two rankings of eighteen values may have nine statistically significant differences in individual placement. But if the nine values, which are the same in both rankings, are the nine most relevant values, the value systems of those informants are much more alike than if the values in the same position are the nine least relevant.

The fourth case is equally susceptible to cultural blindness. It interprets high correlations to indicate considerable sharing of the value system. It interprets the small number of individual differences that detract from the overall high correlations to indicate negligible differences in the systems. But it is possible that just those few values that are different in rank are the central values of the culture. The magnitude of the difference that may exist between those few values can also

not be ignored. Blacks and whites have high correlations between their rankings of values, yet on the issue of equality they differ by ten ranks (Rokeach and Parker 1970, 109). Equality is a key value of our culture, with a high degree of relevance to the history of race relations in the United States. A difference in relative importance as great as ten ranks for a key value may create deep divisions between blacks and whites. The high correlation of their value systems and small number of significant differences might obscure the cleavage.

### ***Three Stages to Understanding: Data Collection, Statistical Analysis, and Interpretation***

I have conceptualized three stages on the path from data to findings in the style of research under consideration (see table 6.6). Each stage may be thought of as self-contained, with principles for adequately applying means to reach its goals. These principles and means have all received a good deal of attention and are relatively well understood. The three stages I have called data collection, analysis, and interpretation.

The goal or output of each stage may be an input for a subsequent stage. Thus data are the output of the data collection stage and are usually the input of the analysis stage. Statistical results are the output of the analysis stage and are usually input for the generation of interpretations and conclusions.

I do not wish to imply that the direction along the path is always from data collection to analysis to interpretation. Expected or hoped-for conclusions and the availability of a particular form of statistical analysis sometimes precede the creation of a measuring instrument.

Table 6.6. Three stages to understanding

Stage	Inputs	Means	Principles	Goals or outputs
Data collection	Phenomena	Measuring instruments	Increase reliability/ decrease experimental bias	Data
SELECTION				
Analysis	Data	Statistics	Applicability of statistical assumptions: normal distribution, interval scale, etc.	Statistical results
EVALUATION OF SIGNIFICANCE				
Interpretation	Statistical results	Reason	Logic	Interpretations and conclusions

I believe that in contrast to the attention that has been given to each of these separate stages in research on reliability and experimental bias (data collection stage), on statistical assumptions and their applicability (analysis stage), and on logic and reasoning (interpretation stage), little attention has been devoted to the transition from one stage to another. We need to study more closely the rules or procedures for passing from data to their statistical analysis, and from statistical results to their interpretation and final presentation as findings. I have shown in one case the many degrees of freedom and the great scope for doubt that exist in the transition from data to analysis and from statistical results to interpretation.

Between the data and the analysis of that data is the uncertainty of selection of a statistical tool and the parts of the data to which it will be applied. Between statistical results and their interpretation is the evaluation of significance. It is replete with uncertainties.

### **The Softening and Silent Ghost of Method Variance**

Hard data have an intrinsic impressiveness which, when combined with a fruitful statistical analysis, may compel the reader to accept the interpretation of the data favored by a scientist. By demonstrating some of the degrees of freedom that exist between data and their statistical analysis, and between the results of the statistical analysis and their interpretation, I hope to have suggested that data are more a method of persuasion in most scholarship and not strictly a mirror of nature or of the truth.

D'Andrade's nonfindings are provocative in a very special way: that the American, Japanese, and Vietnamese subjects in his values questionnaire study are so similar in their reported values may be most telling precisely because, from an ethnographic/thick description point of view, the findings not only defy expectation but also don't really ring true. D'Andrade himself wondered about alternative units for the analysis of values, and it seems fruitful to raise skeptical questions about the soft side of the quantitative analysis of abstract value words and consider disconnects in the general process of moving from cultural reality to tools for measuring, analyzing, and interpreting that reality.

Despite the popularity of cross-cultural questionnaires about abstract value words, the results may be problematic if they fail to attend to concrete institutional realities and to the particular custom complexes of members of local communities. Even if certain formal

standards for questionnaire construction are met, that method for studying cultural values may be somewhat off the mark precisely because the method typically seeks generality of application by removing parochial or culture-specific questions about whether, for example, it is normatively acceptable for widows to eat fish—which are the types of questions that might highlight the cultural differences one is looking for. Whether doing research in Middle Earth or around the corner, collecting, analyzing, and interpreting data are cognitive enterprises involving selection and evaluation. Since the criteria for selection and evaluation are not to be found in the data per se, the data do more to limit the number of credible interpretations of social and cultural reality than to pick out a correct interpretation. Data speak in a very soft voice if they speak at all. They may well say “You have not yet been disproved” but they never boldly assert “You are right!”

## References

- Bourdieu, Pierre. 1972. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- .1990. *A Logic of Practice*. Stanford, CA: Stanford University Press.
- Campbell, Donald T., and Donald W. Fiske. 1959. “Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix.” *Psychological Bulletin* 56 (2): 81–105.
- D’Andrade, Roy G. 2008. *A Study of Personal and Cultural Values*. New York: Palgrave MacMillan.
- Goodnow, Jacqueline J., Peggy J. Miller, and Frank Kessel, eds. 1995. “Cultural Practices as Contexts for Development.” In *New Directions for Child Development*, vol. 67. San Francisco: Jossey-Bass.
- Fassin, Didier. 2012. *A Companion to Moral Anthropology*. Malden, MA: Wiley-Blackwell.
- Hays, William L. 1963. *Statistics for Psychologists*. New York: Holt, Rinehart and Winston.
- Lewin, Kurt. 1943. “Defining the field at given time.” *Psychological Review* 50: 292–310.
- Nagel, Ernest. 1961. *The Structure of Science*. London: Routledge and Kegan Paul.
- Rokeach, Milton, and Seymour Parker. 1970. “Values as Social Indicators of Poverty and Race Relations in America.” In *The Annals of the American Academy of Political and Social Science*, vol. 388, March: 97–111.
- Shweder, Richard A., Manamohan Mahapatra, and Joan G. Miller. 1990. “Culture and Moral Development.” In *Cultural Psychology: Essays on Comparative Human Development*, edited by J. Stigler, R. A. Shweder, and G. Herdt, 1–82. New York: Cambridge University Press.
- Shweder, Richard A., Nancy C. Much, Manamohan Mahapatra, and Lawrence



- Park. 1997. "The Big Three of Morality (Autonomy, Community, Divinity) and The Big Three Explanations of Suffering." In *Morality and Health*, edited by A. Brandt and P. Rozin, 119–72. New York: Routledge and Kegan Paul.
- Weisner, Thomas S. 1984. "A Cross-Cultural Perspective: Ecological Niches of Middle Childhood." In *The Elementary School Years: Understanding Development during Middle Childhood*, edited by A. Collins, 335–69. Washington, DC: National Academy.
- . 1996. "Why Ethnography Should Be the Most Important Method in the Study of Human Development." In *Ethnography and Human Development*, edited by R. Jessor, A. Colby, and R. A. Shweder, 305–26. Chicago: University of Chicago Press.
- . 2001. "Anthropological Aspects of Childhood." In *International Encyclopedia of the Social and Behavioral Sciences* 3: 1697–1701. Oxford: Elsevier Science Ltd.
- . 2002. "Ecocultural Understanding of Children's Developmental Pathways." *Human Development* 45: 275–81.
- Whiting, John W. M., and Irvin Child. 1953. *Child Training and Personality*. New Haven, CT: Yale University Press.