Current Anthropology Volume 40, Supplement, February 1999

2

Culture Consensus as a Statistical Model

by A. Kimball Romney[1]

Two recurring ideas pervade Aunger's paper. The first is that the statistical model chosen by a researcher commits him or her to a philosophical position such as idealism or realism. The second is that a statistical model implies or excludes some explanatory theory. A simple linear regression model describes the relation between two variables according to a linear model; it says nothing about whether the researcher believes the variables are idealistic or realistic, whether they are learned or not, whether they are cultural or physical, and so on. Neither a descriptive scatterplot nor a fitted regression equation implies any kind of causal or other explanatory theory. However, the fit of data to a statistical model may be inconsistent with, and therefore disconfirm, any specific explanatory theory (for an example see Moore and Romney 1994).

Aunger refers to the cultural consensus model, the regression model, and "the bivariate cross-tabulation of societies" as "idealistic" and to "techniques such as log-linear modeling for cross-tabulated data (Agresti 1990), quadratic assignment for similarity matrices (Hubert 1987), and various scaling methods for categorical data (Weller and Romney 1990)" as consistent with his notion of "realism." All of these models (including cultural consensus in Batchelder and Romney 1988) have been proven by the derivation of mathematical theorems to be mathematically consistent. None of them are resistant to inappropriate data (e.g., linear regression will not provide a good fit to strictly nonlinear data); none imply any particular causal model; none have anything to say about whether the researchers or the data are idealist, realist, neither, or both. All of these methods should be included in any scientific tool kit (Romney 1989, 1994).

I happen to believe that culture is both shared and learned. Cultural consensus theory helps describe and measure the extent to which cultural beliefs are shared and in that sense helps me investigate the extent to which my belief is consistent with the data. If the beliefs represented by the data are not shared, the analysis will show this. Data may fit the model and exhibit almost perfect sharing without being cultural or learned. No formal model by itself can indicate whether data are cultural, biological, or physical. Not all sharing is cultural. The formal statistical model of cultural consensus cannot prove that data are cultural (although it can demonstrate that data are not consistent with a shared model) or indicate whether they are learned. In these characteristics the statistical model of cultural consensus is exactly like regression and all other statistical models; it only reveals whether the data are consistent with its mathematical structure.

A Brief Intellectual History

As is generally true of statistical and scientific theories, cultural consensus theory has a deep ancestry. Over 200 years ago Condorcet, a distinguished scientist of the 18th-century Enlightenment, published his jury theorem (1785).[2] Adapting the language of cultural consensus, the jury theorem may be summarized as follows: Given that in a binary-choice situation every person in a group has some minimal (above-chance) "competence" and that individuals choose independently of each other to arrive at a decision on the basis of majority rule, the competence with which the group makes the "correct" decision is greater than individual competence. More important, the collective competence increases with the size of the group and rapidly approaches perfection. This is the same problem as determining, for example, the culturally correct answer to a question like "Is measles contagious?" If there is a correct answer to this question it can be found by independently interviewing a number of appropriate respondents. The title of Batchelder and Romney's (1986) first joint article, "The Statistical Analysis of a General Condorcet Model for Dichotomous Choice Situations," recognizes Condorcet as an intellectual ancestor.

Over 90 years ago Spearman (1904*a, b,* 1910) made fundamental progress in our understanding of reliability and factor analysis, both key concepts in the cultural consensus model (for summary see Romney 1989:174–92). In cultural consensus theory we assume that the correspondence between the answers of any two respondents is a function of the extent to which each is correlated with some truth (Nunnally 1967: chap. 6), an insight originally elucidated by Spearman (1904*b:*255).

---

[1] Department of Anthropology, University of California, Irvine, Calif. 92697-5100, U.S.A. I thank William H. Batchelder, Roy G. D'Andrade, Carmella C. Moore, William C. Watt, and Susan C. Weller for helpful comments on early drafts of this paper. The research was supported in part by National Science Foundation Grant SES-9210009 to A. Kimball Romney and William H. Batchelder.

[2] Ironically, Condorcet, along with the great Lavoisier, was a victim of the French Revolution. As Gillispie wrote, "A tragedy attended Condorcet's faith in science and reason. He wrote in hiding from the guillotine. Nor did he survive the Jacobian Terror of the Revolution, which struck down, too, the scientific institutions of France, partly as survivals of the old regime, partly, too, in a fit of vulgar, sentimental petulance against the hauteur of abstract science, the impersonal tyranny of mathematics, the superiority of the scientist over the artisan" (1960:175).

Approaches used by psychometricians in test construction to study items were adapted to apply to respondents (Lord and Novick 1968, Nunnally 1978). The theory has also been influenced by signal detection theory (Green and Swets 1966). There are structural identities to latent-structure analysis (Lazarsfeld and Henry 1968), although again our applications are to respondents rather than items.

I mention this background to illustrate the way in which normal cumulative progress is made in science. The history also makes explicit the fact that as a statistical model, cultural consensus is imbedded in a rich network of mathematical relationships and structural identities with many other statistical models. Unless one could demonstrate some mathematical error in the derivations of the relevant theorems specific to cultural consensus theory, rejection of the model would require the rejection of latent-structure analysis, psychometric test construction, reliability theory, etc. Yet Aunger's comments do not include any indication that there is anything wrong with our mathematics or with the mathematics of any other statistical model.

When Does Consensus Indicate Cultural Knowledge?

Researchers from a variety of fields have long known that one of the indicators of knowledge is consensus. Given appropriate precautions to guard against bias, collusion, data-collection artifacts, etc., we can probably safely interpret all consensus among respondents as indicating some kind of knowledge. However, it would be a serious error to assume that all consensus indicates *cultural* knowledge.

I think that it is useful to distinguish two broad classes of knowledge. The first arises from the nature of the world and evolutionary processes. Spiders, for example, know how to spin webs, capture prey, etc. Within a given species there will be consensus about the pattern of the web, the mode of capturing prey, etc. I call this kind of knowledge "natural knowledge" to indicate that it is unlearned and arises from interaction between the nature of the organism and that of the organism's world environment. The second kind of knowledge, found mostly in humans, arises from human inventions, is learned and handed down from one generation to the next, and usually varies from one society to another. In cultural consensus theory this kind of knowledge is designated as "cultural knowledge," and language provides the most useful and prototypical example. The line between "natural knowledge" and "cultural knowledge," as between all arbitrarily constructed categories, is not absolute and is sometimes difficult to draw, but the extremes on this continuum and the prototypes are not hard to distinguish at all.

Two necessary if not sufficient parts of the definition of culture are that it is shared among relevant participants and that it is learned as part of our social heritage. The word "relevant" alerts us to the idea that there may be small specialized subgroups, such as medical practitioners, whose members share esoteric knowledge not possessed by the wider cultural group. In short, careful reflection reveals that the very notion of "culture" involves sharing of ideas, concepts, behaviors, etc., by more than one person.

Let us examine an example of cultural sharing discussed by Sapir (1938) more than a half-century ago: the order of the letters in the alphabet. Imagine that we are researchers from outer space and we want to determine whether the order of the alphabet is part of the culture of English-speaking college students in an area called the United States. If we interviewed two students from widely separated areas and observed the same order of 26 letters, the first thing we would conclude is that this consensus was not due to chance ($p = 1/26!$). In order to infer that it was cultural we would have to rule out alternative explanations such as unintended artifacts of the interviewing procedure, prior collusion among the students, answers derivable from the biological and neurological nature of the human species, and answers derivable from the biological and neurological nature of animals in general. Assuming that we could do this for artifacts and collusion, how would we rule out some kind of inborn human universal? A minimum requirement for the confident inference of shared cultural knowledge would be to demonstrate that not all humans could perform the task. For example, one could repeat the task among monolingual Chinese college students who had never seen the English alphabet. If they could not recover the order of the letters of the alphabet we would be more confident about inferring a culturally defined pattern.

To illustrate that not all consensus is cultural I provide some counterexamples. Consider, for example, a task in which we present ten pictures, half of them containing trees, and ask respondents which pictures contain trees. Herrnstein (1979) has performed such an experiment with pigeons and shows that they "can discriminate pictures of trees from pictures lacking trees after minimal training, [and] that the discrimination generalizes to new instances with little or no decrement" (p. 128). This is an example of natural knowledge and would require no learning (beyond the language to understand the questions posed) on the part of human respondents. The fact that this kind of consensus and, thus, knowledge goes beyond humans proves that it cannot be considered cultural.

An example of natural knowledge among humans may be found in the results of an experiment by Boster and Johnson (1989), who presented college students with silhouette line drawings of fish and asked them to sort the drawings in terms of similarity. Presumably any college student anywhere in the world (and any trained monkey) could produce the same general results as found in the article. One would also expect the taxonomic distance to be significantly correlated with the judged similarities as it is. Boster (1987) presents another example of natural knowledge in his paper on judged similarities among bird specimens collected in

South America. He shows high correlations between taxonomic distance and judged similarities, on the one hand, and, on the other, among the judgments of scientists, Aguaruna respondents, Huambisa respondents, and Kentucky college students. He reports that the findings demonstrate that "cultural transmission is apparently not a prerequisite to shared understanding; here, independent observers agree as a result of common inferences from experience" (p. 914).

Returning to the example of the alphabet, it is important to take to heart the observation that whatever "reality" the order of the alphabet possesses resides in the consensus of the people involved. If all the users of the alphabet decided to change the order, they could—as, for example, happened in Arabic. Of course, this would require changing all the alphabet books, encyclopedias, dictionaries, and other printing that depend on the current order. This would be so painful that we cannot conceive of doing it, but in theory it is possible. The bottom line is that the order of the alphabet is coded so well in the minds of all members of the culture that the result is virtually perfect consensus. In this case perfect knowledge is revealed by perfect consensus. The cultural consensus model was developed to investigate culture in situations where consensus is less than perfect.

Scientific Ethics

Over the past few centuries science has developed a well-understood set of conventions and ethical standards. Science is an activity that results in an objective body of knowledge about the world. In normal science this means that new discoveries have to be validated by public replication. Elaborate precautions are taken to ensure that the biases of the investigator do not affect the results. Some of the codification for the collection and analysis of anthropological data may be found in Weller and Romney (1988, 1990).

One convention is reporting an investigation in enough detail that it could be replicated by any competent investigator. This is why, in the empirical examples I discuss below, Weller specified how she obtained the names of the diseases with the free-listing task. This is why she recorded the data in a notebook as they were collected so that she was able to fax me the copies of the originally collected data for me to present below. This is why in the *empacho* paper (Weller et al. 1993) the questions were provided along with summary data. This is why in the original consensus article we published the trivia questions (Romney, Weller, and Batchelder 1986).

Statistical models such as consensus, regression, latent-structure analysis, quadratic assignment, and log-linear modeling make certain assumptions about the kinds of data to which they are appropriately applied. It is an assumption of scientific practice that the investigator will be careful to collect appropriate data. All models make various assumptions about sampling procedures. There are taboos against any kind of data mining and any selection procedures that bias the data. In plain language, it is recognized that the results can be no better than the original data: garbage in, garbage out.

Aunger seems oblivious of these conventions. His description of his research on cultural consensus in the Ituri is so vague that one cannot evaluate it. He says,

> In essence, I used a variety of criteria [what criteria?] to select subsamples [how many?] from the available data. Such manipulations represent reasonably "natural" variation in the database that might be obtained—given the vagaries of ethnographic fieldwork—and hence are legitimate tests of the method's ability to produce the same consensus consistently for a cultural group.
>
> To select data subsamples, I first made assumptions about the nature of the data available for testing (e.g., I changed sample composition by removing all the youngest informants). Second, I simulated other decisions required to formulate an empirical test (e.g., by selecting particular sets of questions which might be more representative of domain knowledge, or by changing the coding of responses).

Such treatment of data is not acceptable in the biomedical research world, nor should it be in the anthropological.

Cultural Consensus Analysis as a Statistical Model

Detailed characteristics of statistical models and scientific theories as generally understood in the natural sciences are discussed elsewhere (Romney 1989, 1994). In this section I sketch some simple examples. A fully abstract theory or model is generally characterized by a set of abstract definitions and axioms together with a number of propositions derived from the axioms. These axioms and propositions form a mathematical model devoid of any real-world content but subject to all the consistency rules of logic. Of equal importance is a set of assumptions and conventions about how empirical variables are fit by the "parameters" in the model. In the simplest cases this enables one to summarize a large amount of data with a very small number of descriptive statistics. In more complicated cases one can predict consequences in the real world from manipulations of the abstract model. The adequacy of the model is judged by how well its predicted consequences correspond to the results of experiments or observations made in the real world.

An example of a simple statistical model is the widespread custom of summarizing data by reporting a mean and standard deviation of a set of measurements from a sample of objects. It is a standard statistical convention that, unless otherwise specified, one is referring to the normal distribution, which is characterized by a mathematical formula with two statistical parameters, a mean and a standard deviation. It is also assumed that a single coherent sample is involved and that the mea-

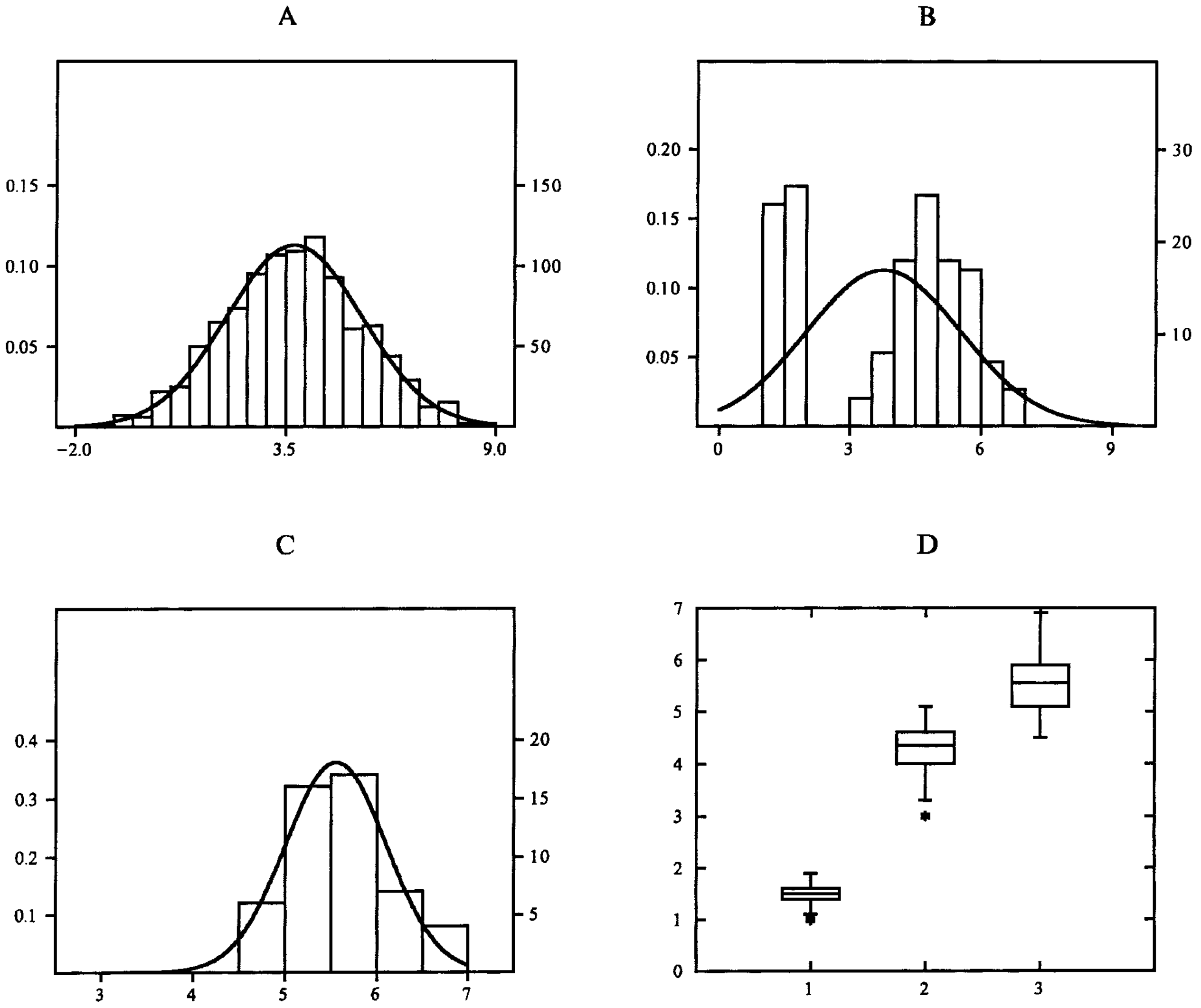

FIG. 1. *Uses of the normal distribution.*

surements involved are appropriate and subject to "error." For example, if I measured petal length of a large sample of iris flowers and reported a mean of 3.758 and a standard deviation of 1.765, a reader would expect the distribution to look something like that in figure 1, *A*. The data for this figure were obtained by taking 1,000 random samples from a normal distribution with the parameters just noted. Unless an empirically obtained sample has roughly this shape, the use of the mean and standard deviation are not appropriate, as they are misleading about the shape of the distribution. This is not because the normal distribution is not useful but rather because these particular data are not appropriately described by it.

Figure 1, *B*, shows an empirically obtained sample of measurements, originally from Fisher, used in SYSTAT to illustrate an inappropriate use of the mean and standard deviation to describe a set of data (Wilkinson 1990). I have fit an idealized mathematical normal distribution to the data with the appropriate parameter values. Wilkinson reports, "We forgot to notice that the petal length measurements involve three different flower species. You can see one of them at the left. The other two are blended at the right. Computing a mean and standard deviation on the mixed data is misleading" (Wilkinson 1990:482). In order to illustrate that the curve fits the individual species I have plotted one of the species in figure 1, *C*, together with the theoretical curve for visual comparison. It would also be possible to make a formal test of fit for each of the species distributions. Finally, I have made a box plot of the three species in figure 1, *D*. Other, more sophisticated models could be applied to the same data, for example, analysis of variance or discriminant analysis.

An anthropological example of a fully developed mathematical process model is cultural consensus for the case of true-false questions (Romney, Weller, and Batchelder 1986; Batchelder and Romney 1986, 1988). This model is easily generalized to multiple-choice and fill-in-the-blank formats. It is characterized by three axioms (the formal mathematical derivations may be found in Batchelder and Romney 1988). In the dichotomous case, the cultural consensus model is mathematically "isomorphic to the two-class latent structure model with the role of respondents and items interchanged" (p. 75). A more informal presentation is found in Romney, Weller, and Batchelder (1986), where the axioms are presented as assumptions:

Assumption 1: Common truth. There is a single answer key shared by all respondents. It is understood that the questions all pertain to a coherent cultural domain and the data are collected in an appropriate format. Recent work has extended the model to accommodate the analysis of intercultural phenomena where the sample is composed of respondents belonging to more than a single culture (Boyd and Johnson n.d., Romney, Moore, and Rusch 1997) or of complicated intracultural variations (Batchelder, Kumbasar, and Boyd 1997).

Assumption 2: Local independence. The respondent-item-response random variables satisfy conditional independence—that is, each respondent's answers are given independently of each other respondent's. It follows that the magnitude of the associations among respondents' answer patterns is a function of the extent to which each respondent is correlated with the answer key, which in anthropological research is usually unknown beforehand.

Assumption 3: Homogeneity of items. Each respondent has a fixed "cultural competence" over all questions. Cultural competence is defined as the proportion of the cultural questions for which the correct answer is known by the respondent. This is a strong assumption that says that questions are all of the same level of difficulty. In some situations one might want to make a weaker assumption, namely, that the respondents who do better on one subset of questions will do better on another subset of questions. This generalization might be called the monotonicity assumption and is related to ensuring that the questions are drawn from a coherent domain. Thus, for example, if tennis experts do better than nonexperts on one subset of the questions concerning tennis, they should do better on another subset.

These assumptions define the ground rules for the operation of the model. They also make it possible to make formal derivations in mathematical terms. Obviously, not all response profile data will conform to these assumptions (Romney, Weller, and Batchelder 1986).

In order to illustrate how the model estimates the cultural competence of each respondent I constructed the artificial and error-free data in figure 2. The five simulated respondents are assumed to have cultural competences represented by the numbers in the shaded right and bottom marginals of the figure. The numbers inside the figure were obtained by multiplying the corresponding row and column marginals (e.g., the second cell in the top row, .72, was obtained by multiplying .90 (the row marginal) by .80 (the column marginal). The multiplication models the consequences of the assumptions above, namely, that the association between any two respondents is a function of the association of each with the "cultural truth" as represented in a shared answer key. Thus, multiplication produces the numbers in the unshaded part of the figure given the numbers in the margins. There is a reverse process called minres (for minimum-residual) factoring (Comrey 1962) that estimates the numbers in the marginals from the numbers in the unshaded part of the figure. In actual applications, the empirical numbers produced by research are the internal unshaded numbers, and we seek to estimate the cultural knowledge of each respondent from them.

	.72	.63	.54	.45	.90	A
.72		.56	.48	.40	.80	B
.63	.56		.42	.35	.70	C
.54	.48	.42		.30	.60	D
.45	.40	.35	.30		.50	E
.90	.80	.70	.60	.50		
A	B	C	D	E		

FIG. 2. *The relation, in error-free hypothetical data, between respondent competences and the magnitude of associations between respondents.*

The index of association that we obtain in an actual study varies with the format of the questionnaire used and is specified by the theory. Formats for which a formal process model has been derived include true-false, multiple-choice, and fill-in-the-blank. There is a "match" method for all three formats in which the observed proportion of matches is corrected for guessing depending upon the number of alternative responses to the question. It is assumed that a respondent answering a fill-in-the-blank question is unlikely to be able to guess the answer, and hence there is no correction for guessing in this case. In addition, there is a "covariance" method for true-false data (described below).

Further insight into figure 2 may be gained by assuming that the internal numbers represent responses to a fill-in-the-blank format in an experiment similar to Boster's (1986) manioc study. In that study he showed each respondent a series of manioc plants and asked for the appropriate name for each. He recorded these names and calculated, for each pair of respondents, the proportion of matches—instances in which the two respondents gave the same name for the plant. These proportions are represented in our hypothetical data by the unshaded areas of figure 2. Given these data, estimating the cultural competence of each respondent with minres factoring is straightforward (Comrey 1962). Another useful characteristic of observed associations is that the square root of the mean value is an approximate estimate of the mean competence of the respondents. The mean of the unshaded figures is .485; the square root of .485 is about .696, a reasonable approximation of the actual mean of .700 of the competences in the shaded marginals.

In a cultural task like knowing the names of fruits and vegetables there is an enormous amount of cultural sharing. For these kinds of data the simplified model illustrated in the assumptions and figure 2 works remarkably well. However, there is nothing in the method that constrains the data to any given form, any more than the existence of a mathematical normal curve constrains the measurements obtained on the length of iris petals. If each person were to give idiosyncratic and random answers to the questions, the data would not fit the model; if we insisted on carrying out the calculations, the average estimated cultural competence of the respondents would be within sampling variability of zero. There are a variety of tests, as in the case of the normal distribution, that allow one to test the extent to which the data fit the model.

Several important consequences have been derived (Batchelder and Romney 1988) from the assumptions of the cultural consensus model for true-false questions. The more important of these are as follows: (1) The knowledge of each of the respondents can be estimated using a ''matching'' method. In this approach respondent competences are computed from a pairwise-agreement matrix among respondents of a match coefficient corrected for guessing. This estimate is invariant under changes in the proportion of items that are true but is affected by bias (the tendency of respondents to say ''true'' when they do not know the answer). (2) The knowledge of each of the respondents can be estimated using a ''covariance'' method. In this approach respondent competences are computed from a pairwise-agreement matrix among respondents of a covariance coefficient corrected for proportion answered ''true.'' This estimate is invariant under changes in bias but is affected by the proportion of the items that are true. (3) By comparing the results of the two methods of estimating knowledge it is possible to determine whether bias or the proportion of questions that are true is unduly affecting the results. (4) An estimate of the ''correct'' answer to each question can be calculated together with a confidence level for classifying the correct answer as ''true'' or ''false.'' (5) How well the data are accounted for by the model can be estimated by examining the eigenvalues (the first should be several times as large as the second and the second only slightly larger than the third) and by ascertaining whether any of the estimates of knowledge are negative (true negative knowledge violates the assumptions of the model). (6) The number of respondents needed to classify a given proportion of the questions correctly with specified confidence levels can be estimated for each average level of knowledge. (7) The expected variance of the distribution of knowledge on the assumption that all respondents have the same knowledge can be calculated. By comparing the observed variance with the expected variance it is possible to determine whether there is any individual variation. As we have explained elsewhere (Romney, Batchelder, and Weller 1987:163–64),

> The cultural consensus model provides a way to utilize much of the accumulated knowledge of traditional psychometric test theory without knowing the ''correct'' answers in advance. The potential implications of this fact could have some profound effects on anthropology and anthropological theory. It means that we are now in a position to measure the knowledge and abilities of informants with a degree of accuracy comparable to that obtained with traditional test theory. This is possible even though test theory depends upon knowledge of the correct answers while consensus theory does not.

Cultural Consensus Analysis of Empirical Data

Analysis by the cultural consensus model of data on the cultural beliefs of 24 urban Guatemalan women about whether each of 27 diseases is contagious or not will help to illustrate that the results are not produced by the model and to demonstrate that concepts such as idealism and realism are inappropriate characterizations of either the data or the model.

The data were provided by Weller and are part of a body of data previously analyzed in two papers written prior to the formalization of consensus theory (Weller 1983, 1984*b*). These data were also presented in the original anthropological description of cultural consensus theory (Romney, Weller, and Batchelder 1986:327–29). Table 1 presents true-false response data for 24 Guatemalan women on judgments about whether each of 27 diseases is contagious or not. ''To ensure that culturally relevant items would be used, [a separate group of] 20 women . . . were asked to name all the illnesses they could think of and to describe each'' (Weller 1984*b:* 342). Of the 27 most frequently listed items, each occurred on at least 15% of the women's lists. In the nearly 20 years since these data were collected I have not seen a more objective and bias-free way of obtaining a sample list of diseases for study. In the original study, building on the work of previous researchers, Weller was investigating whether rank-order judgments of degree of contagion would map in a regular way to a multidimensional scaling representation of judged similarity among diseases (see Weller 1994*b:* fig. 4). There is no way that she could have biased her data in the direction of the consensus model, since the model had not yet been invented, and it was implicitly accepted at the time that beliefs about whether diseases were contagious were a legitimate part of culture to study.

When we apply the match method to the data in table 1 they appear to fit the model quite well. One criterion of a good fit is that the first eigenvalue is very large relative to the second and third. In this example the first

TABLE 1

Dichotomous Data from 24 Urban Guatemalan Women (rows) on Judgments about Whether Each of 27 Diseases (columns) Is Contagious (0) or Noncontagious (1)

0	0	0	0	1	1	1	1	0	1	1	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	0	0	1	1	0	0	1	0	0	0	0	0
0	0	1	1	1	1	1	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
1	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0	1	0	0	0	0	0
1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	1	0	0	0	0
1	0	1	1	1	0	1	1	1	0	1	1	0	0	1	0	0	1	1	0	0	1	1	0	0	0	0
1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0
1	0	0	1	1	0	1	1	1	0	1	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0
1	0	0	1	1	1	1	1	0	0	1	1	0	1	0	1	0	1	1	0	0	1	0	0	0	0	0
1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	0	1	1	1	1	1	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	1	0	1	0	0	0	1	0
1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	1	0	1	0	0	0	1	0
1	0	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	0	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0
1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0	1	1	0	0	0	0
1	0	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	0	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0
1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0
1	0	0	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	0	0	1	0	0	0	0	0

three eigenvalues are 16.363, 1.580, and 0.933, giving a ratio of 10.358 between the first and second and 1.694 between the second and third. All 24 respondents have positive competence scores, with a mean of 0.819 and a standard deviation of 0.104. The application of the covariance method (assuming a π of 0.5) gives equally good results, with the first three eigenvalues of 16.929, 1.243, and 0.945 producing a ratio of 13.616 between the first and second and 1.303 between the second and third. All 24 respondents have positive competence scores, with a mean of 0.834 and a standard deviation of 0.102. The correlation between the two sets of competence scores is $r = 0.994$. Since the match method is sensitive to bias and the covariance method is not, this high correlation indicates that bias is not a significant factor in these data. For both methods all diseases are significantly classified as either contagious or noncontagious, and this answer key calculated with Bayesian methods corresponds with the modal, or most frequent, response. This outcome is not unusual for high-consensus data and large sample sizes (e.g., 20 or more).

There is a very nice fit between the data and the model. The model provides an estimate of the cultural competence of each respondent as well as a classification of the diseases into contagious and noncontagious ones. The estimates of cultural competence, like all statistical estimates, suggest further questions relating to why some women "know more" than others about the contagiousness of diseases. We are in a position to test ideas about whether age or experience, as indexed by number of children, might relate to cultural knowledge. Table 2 shows the competence estimates (computed by both covariance and match methods) for the 24 respondents together with the age and the number of children

TABLE 2

Covariance and Match Competence Scores on Contagion, Age, and Number of Children for 24 Guatemalan Respondents

ID	Covariance Competence	Match Competence	Age	Number of Children
1	0.985	0.967	26	3
2	0.985	0.967	28	4
3	0.948	0.936	29	1
4	0.948	0.936	36	3
5	0.947	0.923	26	4
6	0.925	0.918	48	4
7	0.927	0.909	38	1
8	0.910	0.893	25	1
9	0.884	0.867	50	6
10	0.880	0.863	44	2
11	0.846	0.845	28	2
12	0.831	0.820	52	3
13	0.817	0.807	27	3
14	0.839	0.804	26	1
15	0.790	0.790	29	2
16	0.787	0.786	32	2
17	0.774	0.775	33	2
18	0.772	0.757	22	2
19	0.738	0.742	28	1
20	0.760	0.740	70	4
21	0.759	0.726	26	1
22	0.757	0.723	31	2
23	0.616	0.592	28	2
24	0.610	0.574	26	2

of each. When we compare the cultural knowledge of the women under 30 with that of the women over 30, we find no significant differences. For the covariance method the mean competence of the 14 women under 30 years of age is 0.826, while the mean competence of the 10 women over 30 years of age is 0.847. The difference tested with an independent-samples *t* test is not significant. The situation is different when experience is indexed by number of children. Here, the 15 women with only one or two children have a mean competence computed with the covariance method of 0.797, while the 9 women with three or more children have a mean competence of 0.898, a 0.101 difference that is significant at the 0.02 level with an independent-samples *t* test (the results are virtually identical on the match-derived competences). This result is entirely in line with a commonsense approach to cultural knowledge. Women who have more children have more opportunity to learn about whether a given disease is contagious and therefore of potential danger to their children.

In order to dramatize these results I have arranged the diseases in order of least (top) to most contagious and recorded the number of women agreeing on that response (table 3). On the basis of binomial theory one can assign a probability to the number of 1's or 0's on the assumption of equal chance of either. On this basis all but three of the diseases are classified as either contagious or noncontagious beyond a chance level. The space in the middle of the table separates the contagious from the noncontagious simply on the basis of the most frequent or modal response. The women are arranged in approximate order—most agreement on the left, least on the right (mostly true whether based on agreement with modal response or on patterns of sharing). The first two women agree perfectly with each other on all 27 diseases (this would occur by chance once in 134, 217, 728 trials), as do the second two. There are six diseases never judged contagious and three diseases always judged contagious. One can tell by casual inspection that something other than chance has been involved in producing these data.

In order to demonstrate that not all data fit the cultural consensus model, I now turn to some additional data collected by Weller at the same time as those reported above. What is most interesting about this data set is that the study was originally designed to show how the hot-cold concept would apply to diseases in Guatemala and Mexico. In table 4 the diseases have

TABLE 3
Dichotomous Response Data for 24 Guatemalan Women Arranged from Most to Least Competent on Contagiousness of 27 Diseases (1, Noncontagious; 0, Contagious) Arranged from Least to Most Contagious

	Respondent																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Arthritis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Colic	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Diabetes	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Kidney pain	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Gastritis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Rheumatism	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24**
Appendicitis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	23**
Cancer	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	22**
Intestinal influenza	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	22**
Tetanus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	21**
Allergies	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	20**
Diarrhea	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	20**
Polio	1	1	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1	0	0	0	1	17
Malaria	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	0	1	1	0	0	0	0	16
Tonsilitis	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0	1	0	1	0	1	1	1	0	0	10
Amoebas	0	0	1	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0	0	0	8*
Diphtheria	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	5**
Hepatitis	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	3**
Typhoid fever	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	3**
Chicken pox	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2**
Rubella	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2**
Flu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1**
Mumps	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1**
Tuberculosis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1**
Whooping cough	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0**
Smallpox	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0**
Measles	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0**
Errors	0	0	1	1	1	2	2	2	2	2	3	3	3	3	3	3	3	4	3	4	5	5	6	6	

**$p < .01$, *$p < .05$ (binomial test)

TABLE 4

True/False Response Data for 23[a] *Guatemalan Women Arranged from Most to Least Competent on "Hot/Cold Remedies" Needed for 27 Diseases (1, Needs Hot Remedy; 0, Needs Cold Remedy) Arranged from Most Hot to Most Cold*

	Respondent																							
	10	17	9	23	22	6	18	2	3	5	8	21	1	14	7	13	16	4	20	12	19	15	24	
Allergies	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	20**
Kidney pain	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	0	19**
Gastritis	1	1	1	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	0	1	0	1	16*
Amoebas	1	0	1	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	1	1	1	15
Appendicitis	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	0	1	1	0	1	0	0	15
Hepatitis	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	14
Mumps	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	13
Rubella	1	1	0	1	1	0	1	0	1	0	0	0	1	1	0	0	1	1	1	1	1	0	0	13
Measles	1	1	0	1	1	0	1	0	1	1	0	0	1	1	0	0	0	1	1	1	1	0	0	13
Smallpox	1	1	1	1	1	0	1	0	1	1	0	0	1	1	0	0	1	0	0	1	1	0	0	13
Cancer	1	1	1	0	0	1	0	1	0	1	1	1	0	0	1	0	0	0	1	0	0	1	1	12
Diabetes	1	1	1	1	0	1	1	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	0	12
Intestinal influenza	1	0	1	1	1	1	1	1	0	1	1	0	0	0	1	1	0	0	0	0	0	1	0	12
Tetanus	1	1	1	1	0	1	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0	1	12
Chicken pox	1	1	0	1	1	0	1	0	1	0	0	0	1	1	0	0	1	1	0	1	1	0	0	12
Tonsilitis	0	0	1	1	0	0	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	0	0	11
Polio	0	0	1	0	1	1	0	1	1	1	1	1	0	0	1	0	0	0	1	0	0	0	1	11
Diarrhea	0	0	1	1	0	1	1	1	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	10
Typhoid fever	0	1	0	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	10
Diphtheria	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	1	7*
Arthritis	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	6**
Whooping cough	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	6**
Tuberculosis	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	1	1	6**
Malaria	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	1	0	0	5**
Colic	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	4**
Rheumatism	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	3**
Flu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1**
Errors	2	3	7	6	4	7	4	10	6	6	9	8	4	6	8	8	6	9	9	6	7	12	12	

$^{**}p < .01$, $^{*}p < .05$ (binomial test)

[a] Respondent 11 is missing; otherwise, respondent number is the same as for the contagion data.

been arranged from those requiring a "hot" remedy (top) to those requiring a "cold" remedy (bottom) in terms of the number of women agreeing on the response. The amount of agreement among respondents is much less than in the case of contagion. The cultural consensus model does not fit well, and the first factor does not represent competence. What little competence is present is found in the second factor. Only a minority of the diseases are classified into one or the other category by any criterion including the binomial. No women have the same profile of answers as was the case for contagion. There is some agreement on a subset of the diseases; for example, the women seem to agree that allergies, kidney pain, and gastritis take "hot" remedies while diphtheria, arthritis, whooping cough, tuberculosis, malaria, colic, rheumatism, and flu take "cold" remedies. By and large, however, the two tables differ so much that we would characterize the two data sets as qualitatively different. The hot-cold concept does not exhibit shared knowledge across all diseases in the same sense as does contagion and noncontagion. Even though there may be beliefs about a specific subset of illnesses, the overall cultural pattern is clearly different.

The utility of cultural consensus theory in comparative research is illustrated in a recent study of knowledge and beliefs about the disease *empacho,* a gastrointestinal disorder (Weller et al. 1993). The researchers had each done previous work on *empacho* in widely dispersed areas and in this study collaborated using a single method that allowed comparisons not only across sites but with previously published results. The sites studied were a rural town in northeastern Guatemala, the urban town of Guadalajara, Mexico, with interviews with both urban and rural mestizos, Hidalgo County in South Texas among Mexican-Americans, and a Latino population of Puerto Rican respondents in Hartford, Connecticut. Consensus theory showed some small variations among sites but overall found very large consistency in results across the sites that "suggests a common origin for the concept of empacho" (Weller et al. 1993:122). "A comparative study based upon a standard protocol is one of the most powerful methodological tools there is. . . . [and] consistency in results across four diverse settings leads to the inference that a similar consistency in beliefs about empacho might be found in the encompassed region'' (p. 123).

It should also be reported that consensus theory has been subjected to extensive testing through simulation (Maher 1987, Weller 1987) and Monte Carlo methods. A small sample of situations in which it has been applied would include folk medical beliefs (Garro 1986, 1988; Ruebush, Weller, and Klein 1992; Weller et al. 1993), judgments of personality traits in a college sorority (Iannucci 1991, Iannucci and Romney 1994), semiotic characterizations of alphabetic systems (Jameson 1989, Jameson and Romney 1990), occupational prestige (Romney 1989), causes of death (Romney, Batchelder, and Weller 1987), illness beliefs of deaf senior citizens (Steinhaus-Donham 1987), hot-cold concepts of illness (Weller 1983, 1984*b*), child abuse (Weller, Romney, and Orr 1986), graffiti writers' evaluations of strategies to control illegal graffiti (Brewer 1992), and national consciousness in Japan (Yoshino 1989). The validity of the theory is also much enhanced by the fact that cultural competence has unanticipated associations with other social and psychological characteristics, as pointed out in an important contribution by D'Andrade (1987). The characteristics discussed by D'Andrade include reliability, consistency, normality, education, intelligence, and experience.

Have we learned anything definite through the use of systematic data collection methods and cultural consensus analysis that we would not otherwise know? I think the answer is clearly yes. First, the use of the free-listing task is a standard, objective way to obtain a meaningful sample of the domain under investigation. My colleagues and I have helped develop, describe, and use this method over many years (e.g., Romney and D'Andrade 1964, Weller and Romney 1988, Romney, Moore, and Rusch 1997). It is as nearly free of investigator bias as any method invented. It allows a separate investigator to return to the same area and obtain an independent list of diseases that will be very similar to the original regardless of who does the collecting and with what political or other biases.

Second, by asking a sample of women exactly the same question about every disease we were able to obtain objective information about their beliefs. On average the women knew just over 80% of the correct and hence, for them, ''true'' answers to the contagion questions. What could be wrong with asking women whether the diseases their children might encounter are contagious or require a hot or cold remedy? Would I have been less ''idealistic'' had the women answered the questions at random? Aunger seems to think that it is my idealism that led the women to agree with each other. On the contrary, I am merely summarizing the data and conclude that for these women the culturally correct and in this sense ''true'' answer to the question whether measles is contagious is yes. It would be culturally correct (or ''true'') even if factually wrong (i.e., if measles were not contagious).

Third, the women differed slightly in the amount of cultural knowledge they had as estimated by the method. We were able to demonstrate that the women with three or more children knew, on average, about 10% more than the women with only one or two children. There were no significant differences based on age. This kind of precise, testable, and replicable knowledge is simply not obtainable without systematic data collection and the statistical model of cultural consensus.

Fourth, the unexpected and surprising finding that the hot-cold concept does not apply to diseases in the same sense that the concept of contagion does calls into question the findings of traditional ethnographic practices. All previous methods had failed to indicate that anthropologists had been on the wrong track for decades with regard to the concept of hot-cold remedies, and I had even helped in keeping the myth alive (D'Andrade et al. 1972).

Fifth, the findings of the Weller et al. (1993) comparative paper are very important and could not have been obtained without objective and systematic methods of analysis. No one would have predicted that such widely dispersed communities would have such a high degree of sharing of beliefs about a single disease.

To reiterate, none of these findings would have been possible without systematic data collection and appropriate statistical analysis. Aunger does not have an alternative way of getting similar types of answers. His methodological and theoretical comments are not even wrong.[3]

I believe that the belief that a given disease is contagious is a learned cultural response. This belief is neither supported nor disconfirmed by the data. Without their labels there is nothing in the data of table 1 that tells us whether they are cultural or learned. The table could just as well have been produced by a behavioral ecologist studying the foraging dynamics of bumblebees, with the 1's and 0's representing movements within and between plant species (Chittka, Gumbert, and Kunze 1997). Had that been the case, would we want to characterize the data as either idealistic or realistic? I think not. If the data derive from bees we probably don't want to call them cultural (although they may have a learned component). What is invariant about them is just this: there is consensus among the respondents concerning some characteristic of the items.

If the data are not appropriately characterized as idealistic or realistic, then does it make more sense to say that Romney the researcher is one or the other, neither, or both? Do I become an idealist when I borrow Weller's neat and accurate field recordings and put them in table 1? Aunger considers it ''a moral imperative that individ-

---

[3] Anthropologists would do well to ponder the implications of Orans's marvelous book, *Not Even Wrong.* He comments: ''Logically contradictory evidence refutes an argument. Constructing arguments that are capable of refutation is the hallmark of science. I cannot think of a polite term for arguments not admitting of refutation; perhaps the phrase 'not even wrong' is all the condemnation that is required'' (1996:133).

uals be considered real, their minds unique. They should be treated with respect and valued for their diversity of experience and opinion.'' How better to treat individuals with respect than by asking them in as unbiased a manner as possible to indicate their beliefs and then to record their responses with as much objective accuracy and scientific integrity as possible?

Both sharing and learning are important characteristics of culture. It is an interesting thought experiment to try to construct a realistic scenario in which something that has no sharing whatsoever can involve learning. What is there to learn? People share the cultural answer. Without the sharing how do we arrive at an answer? The problem is that Aunger says, ''If a belief is learned from others, then it is cultural; if it is invented or inferred from individual experience, it is not.'' I agree. However, it seems to me obvious that if there is no sharing one cannot discern what is being learned from others. In the data above it works perfectly well to assume that the modal response defines what the majority of the sample of the culture think the answer is. That is certainly a better approximation than what the anthropologist thinks after talking to one or two respondents. Cultural consensus provides the best estimate to date of the cultural answer to cultural questions.

Wider Implications

Aunger's negative attitude concerning the aims of traditional scholarship and science is a symptom of only one of many malaises characterizing social anthropology today. As a result, social anthropology is very nearly moribund as a field contributing to a scientific knowledge of human behavior. A few years ago, in a short sketch of my intellectual career, I posed the question ''What are the prospects that social anthropology will ever become a science and accumulate knowledge of the cultural aspects of human behavior?'' (Romney 1994:276). When I embarked upon my career at midcentury the prospects seemed bright. Because of vast changes in attitudes and beliefs in both universities and the wider culture, I suspect that now the chances are very close to zero. Social anthropology will never attain the status of a mature science that accumulates knowledge if it does not aspire to become a science at all. Even were there to be a shift in the current secular trend, knowledge and appreciation of statistical methods, including consensus theory, are virtually nonexistent. The will, the skill, and the modesty required to bring our ideas into the ''empirical arena'' where each assertion has to be tested by the appropriate ''objective'' methodology are lacking.

References Cited

AGRESTI, A. 1990. *Categorical data analysis.* New York: John Wiley.

ATRAN, SCOTT. 1990. *Cognitive foundations of natural history.* Cambridge: Cambridge University Press.

AUNGER, ROBERT. 1995. On ethnography: Story-telling or science? CURRENT ANTHROPOLOGY 36:97–130.

———. n.d. *a.* No characterization of cultural consensus or individual competence is definitive. MS.

———. n.d. *b.* The cultural transmission of belief in a face-to-face society. MS.

———. n.d. *c.* Exposure versus susceptibility in the epidemiology of ''everyday'' beliefs. MS.

BARTH, FREDERIK. 1987. *Cosmologies in the making: A generative approach to cultural variation in Inner New Guinea.* Cambridge: Cambridge University Press.

———. 1993. *Balinese worlds.* Chicago: University of Chicago Press.

BATCHELDER, WILLIAM H., ECE KUMBASAR, AND JOHN P. BOYD. 1997. Consensus analysis of three-way social network data. *Journal of Mathematical Sociology* 22:29–58.

BATCHELDER, WILLIAM H., AND A. KIMBALL ROMNEY. 1986. ''The statistical analysis of a general Condorcet model for dichotomous choice situations,'' in *Information pooling and group decision making.* Edited by B. Grofman and G. Owen, pp. 103–12. Greenwich, Conn.: JAI Press.

———. 1988. Test theory without an answer key. *Psychometrika* 53:71–92.

BATCHELDER, WILLIAM H., AND A. KIMBALL ROMNEY. 1989. ''New results in test theory without an answer key,'' in *Mathematical psychology in progress.* Edited by E. E. Roskam, pp. 229–48. Heidelberg: Springer-Verlag.

BIDNEY, DAVID. 1967. 2d edition. *Theoretical anthropology.* New York: Schocken Books.

BORGATTI, STEPHEN P. 1994. Cultural domain theory. *Journal of Quantitative Anthropology* 4:261–78.

BOROFSKY, ROBERT. 1994. ''Introduction,'' in *Assessing cultural anthropology.* Edited by Robert Borofsky. New York: McGraw-Hill.

BOSTER, JAMES S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88:429–36.

———. 1987. Agreement between biological classification systems is not dependent on cultural transmission. *American Anthropologist* 89:914–19.

———. 1991. ''The information economy model applied to biological similarity judgement,'' in *Socially shared cognition.* Edited by J. Levine, L. Resnick, and S. Behrend, pp. 203–25. Washington, D.C.: American Psychological Association.

BOSTER, JAMES S., AND ROY G. D'ANDRADE. 1989. Natural and human sources of cross-cultural agreement in ornithological classification. *American Anthropologist* 91:132–42.

BOSTER, JAMES S., AND JEFFREY C. JOHNSON. 1989. Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist* 91:866–89.

BOSTER, JAMES S., AND SUSAN C. WELLER. 1990. Cognitive and contextual variation in hot-cold classification. *American Anthropologist* 92:171–78.

BOYD, JOHN P., AND JEFFREY C. JOHNSON. N.D. Finding *N*-culture consensus with simulated annealing. MS.

BOYD, ROBERT, AND PETER J. RICHERSON. 1985. *Culture and the evolutionary process.* Chicago: University of Chicago Press.

BOYER, PASCAL. 1994. *The naturalness of religious ideas: A cognitive theory of religion.* Berkeley: University of California Press.

BREWER, DEVON A. 1992. Hip hop graffiti writers' evaluations of strategies to control illegal graffiti. *Human Organization* 51:188–96.

———. 1995. Cognitive indicators of knowledge in semantic domains. *Journal of Quantitative Anthropology* 5:107–28.

CARSPECKEN, PHILIP F., AND M. APPLE. 1992. ''Critical research: Theory, methodology, and practice,'' in *The handbook of qualitative research in education.* Edited by M. D. LeCompte, W. L. Millroy, and J. Preissle, pp. 507–54. New York: Academic Press.

CAVALLI-SFORZA, LUIGI L., AND MARCUS W. FELDMAN.

1981. *Cultural transmission and evolution: A quantitative approach.* Princeton: Princeton University Press.

Chavez, Leo R., F. Allan Hubbell, Juliet M. McMullin, Rebecca G. Martinez, and Shiraz I. Mishra. 1995. Structure and meaning in models of breast and cervical-cancer risk-factors: A comparison of perceptions among Latinas, Anglo women, and physicians. *Medical Anthropology Quarterly* 9:40–74.

Chittka, Lars, Andreas Gumbert, and Jan Kunze. 1997. Foraging dynamics of bumble bees: Correlates of movements within and between plant species. *Behavioral Ecology* 8: 239–49.

Clifford, James. 1988. *The predicament of culture: Twentieth-century ethnography, literature, and art.* Cambridge: Harvard University Press.

Clifford, James, and George E. Marcus. Editors. 1986. *Writing culture: The poetics and politics of ethnography.* Berkeley: University of California Press.

Comrey, Andrew L. 1962. The minimum residual method of factor analysis. *Psychological Reports* 11:15–18.

Condorcet, Nicolas Caritat de. 1785. *Essai sur l'application de l'analyse a la probabilité des décisions rendues à la pluralité des voix.* Paris.

D'Andrade, Roy G. 1987. Modal responses and cultural expertise. *American Behavioral Scientist* 31:194–202.

———. 1989*a*. "Cultural cognition," in *Foundations of cognitive science.* Edited by Michael I. Posner, pp. 795–830. Cambridge: MIT Press.

———. 1989*b*. "Some propositions about the relations between culture and human cognition," in *Cultural psychology: Essays on comparative human development.* Edited by James W. Stigler, Richard A. Shweder, and Gilbert Herdt, pp. 65–129. Cambridge: Cambridge University Press.

———. 1995. *The development of cognitive anthropology.* Cambridge: Cambridge University Press.

D'Andrade, Roy G., Naomi Quinn, Sara B. Nerlove, and A. Kimball Romney. 1972. "Categories of disease in American English and Mexican Spanish," *Multidimensional scaling: Theory and application in the behavioral sciences,* vol. 2, *Applications.* Edited by A. Kimball Romney, Roger Shepard, and Sara B. Nerlove, pp. 11–54. New York: Seminar Press.

Ford, Clellan S. Editor. 1967. *Cross-cultural approaches: Readings in comparative research.* New Haven: HRAF Press.

Garro, Linda. 1986. Intracultural variation in folk medical knowledge: A comparison between groups. *American Anthropologist* 88:351–70.

———. 1988. Explaining high blood pressure: Variation in knowledge about illness. *American Ethnologist* 15:98–119.

———. n.d. Intracultural variation in causal accounts of diabetes: A comparison of three Canadian Anishinaabe (Ojibway) communities. *Culture, Medicine, and Psychiatry.* In press.

Geertz, Clifford. 1988. *Works and lives: The anthropologist as author.* Stanford: Stanford University Press.

Gillispie, Charles C. 1960. *The edge of objectivity: An essay in the history of scientific ideas.* Princeton: Princeton University Press.

Green, David M., and John A. Swets. 1966. *Signal detection theory and psychophysics.* New York: John Wiley.

Guglielmino, C. R., C. Viganotti, Barry Hewlett, and Luigi L. Cavalli-Sforza. 1995. Cultural variation in Africa: The role of mechanisms of transmission and adaptation. *Proceedings of the National Academy of Sciences, U.S.A.* 92:7585–89.

Herrnstein, Richard J. 1979. Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes* 5:116–29.

Hewlett, Barry S., and Luigi L. Cavalli-Sforza. 1986. Cultural transmission among the Aka pygmies. *American Anthropologist* 88:922–33.

Hirschfeld, Lawrence A., and Susan A. Gelman. Editors. 1994. *Mapping the mind: Domain specificity in cognition and culture.* Cambridge: Cambridge University Press.

Hubert, Lawrence J. 1987. *Assignment methods in combinatorial data analysis.* New York: Marcel Dekker.

Iannucci, Anita L. 1991. Patterns of consensus and divergence in the judgment of personality traits among a group of well-acquainted young women. Ph.D. diss., University of California, Irvine, Calif.

Iannucci, Anita L., and A. Kimball Romney. 1994. Consensus in the judgment of personality traits among friends and acquaintances. *Journal of Quantitative Anthropology* 4: 279–95.

Jameson, Kimberly. 1989. An empirical investigation of semiotic characterizations of alphabetic systems. Ph.D. diss., University of California, Irvine, Calif.

Jameson, Kimberly, and A. Kimball Romney. 1990. Consensus on semiotic models of alphabetic systems. *Journal of Quantitative Anthropology* 2:289–303.

Johnson, Jeffrey C., and David C. Griffith. 1996. Pollution, food safety, and the distribution of knowledge. *Human Ecology* 24:87–108.

Johnson, Kathy E., Carolyn B. Mervis, and James S. Boster. 1992. Developmental changes within the structure of the mammal domain. *Developmental Psychology* 28: 74–83.

Keesing, Roger M. 1974. Theories of culture. *Annual Review of Anthropology* 3:73–97.

———. 1987. Anthropology as interpretive quest. current anthropology 28:161–76.

Keil, Frank C. 1995. "The growth of causal understandings of natural kinds," in *Causal cognition: A multidisciplinary debate.* Edited by Dan Sperber, David Premack, and Ann James Premack, pp. 234–66. Oxford: Oxford University Press.

Kroeber, A. L., and C. Kluckhohn. 1952. Culture: A critical review of concepts and definitions. *Papers of the Peabody Museum of American Archaeology and Ethnology* 47:1–223.

Lazarsfeld, Paul F., and N. W. Henry. 1968. *Latent structure analysis,* New York: Houghton Mifflin.

Lord, F. M., and M. R. Novick. 1968. *Statistical theories of mental test scores.* Reading, Mass.: Addison-Wesley.

McMullin, Juliet M., Leo R. Chavez, and F. Allan Hubbell. 1996. Knowledge, power, and experience: Variation in physicians' perceptions of breast cancer risk factors. *Medical Anthropology* 16:295–317.

Mack, John E. 1994. *Abduction: Human encounters with aliens.* New York: Ballantine.

Maher, Kathleen M. 1988. A multiple choice model for aggregating group knowledge and estimating individual competencies. Ph.D. diss., University of California, Irvine, Calif.

Marks, N. L., and H. Miller. 1987. Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin* 8:728–35.

Mathews, Holly F. 1983. Context-specific variation in humoral classification. *American Anthropologist* 85:826–43.

Moore, Carmella C., and A. Kimball Romney. 1994. Material culture, geographic propinquity, and linguistic affiliation on the North Coast of New Guinea: A reanalysis of Welsch, Terrell, and Nadolski (1992). *American Anthropologist* 96:370–96.

Nunnally, Jum C. 1967. *Psychometric theory.* New York: McGraw-Hill.

Orans, Martin. 1996. *Not even wrong: Margaret Mead, Derek Freeman, and the Samoans.* Novato, Calif.: Chandler and Sharp.

Roberts, John M. 1964. "The self-management of cultures," in *Explorations in cultural anthropology.* Edited by Ward H. Goodenough, pp. 433–54. New York: McGraw-Hill.

Romney, A. Kimball. 1989. Quantitative models, science, and cumulative knowledge. *Journal of Quantitative Anthropology* 1:153–223.

———. 1994. "Cultural knowledge and cognitive structure," in *The making of psychological anthropology 2.* Edited by Marcelo M. Suarez-Orozco, George Spindler, and Louise Spindler, pp. 254–83. New York: Harcourt Brace.

Romney, A. Kimball, William H. Batchelder, and

SUSAN C. WELLER. 1987. Recent applications of cultural consensus theory. *American Behavioral Scientist* 31:163–77.

ROMNEY, A. KIMBALL, JOHN P. BOYD, CARMELLA C. MOORE, WILLIAM H. BATCHELDER, AND TIMOTHY J. BRAZILL. 1996. Culture as shared cognitive representations. *Proceedings of the National Academy of Sciences, U.S.A.* 93: 4699–4705.

ROMNEY, A. KIMBALL, AND ROY G. D'ANDRADE. 1964. Cognitive aspects of English kinship terms. *American Anthropologist,* 66, pt. 2:146–270.

ROMNEY, A. KIMBALL, CARMELLA C. MOORE, AND CRAIG D. RUSCH. 1997. Cultural universals: Measuring the semantic structure of emotion terms in English and Japanese. *Proceedings of the National Academy of Sciences, U.S.A.* 94: 5489–94.

ROMNEY, A. KIMBALL, SUSAN C. WELLER, AND WILLIAM H. BATCHELDER. 1986. Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist* 88: 313–38.

RUEBUSH, TRENTON K., SUSAN C. WELLER, AND ROBERT E. KLEIN. 1992. Knowledge and beliefs about malaria on the Pacific coastal plain of Guatemala. *American Journal of Tropical Medicine and Hygiene* 46:451–59.

SAPIR, EDWARD. 1938. Why cultural anthropology needs the psychiatrist. *Psychiatry* 1:7–12. (Reprinted in *Selected writings of Edward Sapir,* edited by David G. Mandelbaum, pp. 569–77. Berkeley: University of California Press, 1949.)

SOLTIS, JOSEPH, ROBERT BOYD, AND PETER J. RICHERSON. 1995. Can group-functional behaviors evolve by cultural group selection? An empirical test. CURRENT ANTHROPOLOGY 36:473–94.

SPEARMAN, CHARLES. 1904*a.* The proof and measurement of association between two things. *American Journal of Psychology* 15:720–801.

———. 1904*b.* General intelligence, objectively determined and measured. *American Journal of Psychology* 15:201–93.

———. 1910. Correlation calculated from faulty data. *British Journal of Psychology* 3:271–95.

SPERBER, DAN. 1985. *On anthropological knowledge.* Cambridge: Cambridge University Press.

———. 1991. ''The epidemiology of beliefs,'' in *Psychological studies of widespread beliefs.* Edited by Colin Fraser and George Gaskill, pp. 25–43. Oxford: Oxford University Press.

———. 1994. ''The modularity of thought and the epidemiology of representations,'' in *Mapping the mind: Domain specificity in cognition and culture.* Edited by Lawrence A. Hirschfeld and Susan A. Gelman, pp. 39–67. Cambridge: Cambridge University Press.

———. 1996. *Explaining culture: A naturalistic approach.* Oxford: Blackwell.

SPSS, INC. 1990. *SPSS 4.0.* Chicago.

STEINHAUS-DONHAM, CYNTHIA A. 1987. Illness beliefs of deaf senior citizens. Ph.D. diss., University of California, Irvine, Calif.

STRAUSS, CLAUDIA, AND NAOMI QUINN. 1994. ''A cognitive/cultural anthropology,'' in *Assessing cultural anthropology.* Edited by Robert Borofsky, pp. 284–97. New York: McGraw-Hill.

SWARTZ, MARC J. 1982. Cultural sharing and cultural theory: Some findings of a five-society study. *American Anthropologist* 84:314–38.

———. 1991. *The way the world is: Cultural processes and social relations among the Mombasa Swahili.* Berkeley: University of California Press.

TOMASELLO, MICHAEL, ANN CALE KRUGER, AND HILARY HORN RATNER. 1993. Cultural learning. *Behavioral and Brain Sciences* 16:495–552.

TOOBY, JOHN, AND LEDA COSMIDES. 1992. ''The psychological foundations of culture,'' in *The adapted mind: Evolutionary psychology and the generation of culture.* Edited by Jerome Barkow, Leda Cosmides, and John Tooby, pp. 19–136. New York: Oxford University Press.

VAYDA, ANTHONY P. 1994. ''Actions, variations, and change: The emerging anti-essentialist view in anthropology,'' in *Assessing cultural anthropology.* Edited by Robert Borofsky, pp. 320–29. New York: McGraw-Hill.

WALLACE, ANTHONY F. C. 1970. 2d edition. *Culture and personality.* New York: Random House.

WEBER, MAX. 1971 (1949). ''Ideal types and theory construction,'' in *Readings in the philosophy of the social sciences.* Edited by May Brodbeck, pp. 496–507. London: Macmillan.

WELLER, SUSAN C. 1983. New data on intracultural variability: The hot/cold concept of medicine and illness. *Human Organization* 42:249–57.

———. 1984*a.* Consistency and consensus among informants: Disease concepts in a rural Mexican town. *American Anthropologist* 86:966–75.

———. 1984*b.* Cross-cultural concept of illness: Variation and validation. *American Anthropologist* 86:341–51.

———. 1987. Shared knowledge, intracultural variation, and knowledge aggregation. *American Behavioral Scientist* 31:178–93.

WELLER, SUSAN C., LEE M. PRACHTER, ROBERT T. TROTTER, AND ROBERTA D. BAER. 1993. *Empacho* in four Latino groups: A study of intra- and inter-cultural variation in beliefs. *Medical Anthropology* 15:109–36.

WELLER, SUSAN C., AND A. KIMBALL ROMNEY. 1988. *Systematic data collection.* Newbury Park, Calif.: Sage.

———. 1990. *Metric scaling: Correspondence analysis.* Newbury Park, Calif.: Sage.

WELLER, SUSAN C., A. KIMBALL ROMNEY, AND DONALD P. ORR. 1986. The myth of sub-culture of corporal punishment. *Human Organization* 46:39–47.

WILKINSON, LELAND. 1990. *SYSTAT: The system for statistics.* Evanston, Ill: SYSTAT.

YOSHINO, RYOZO. 1989. An extension of the ''test theory without answer key'' by Batchelder and Romney and its application to an analysis of data of national consciousness (in Japanese). *Proceedings of the Institute of Statistical Mathematics* 37: 171–88.