ANTHROPOLOGICAL RECORDS
1:4

CULTURE ELEMENT DISTRIBUTIONS: VIII
THE RELIABILITY OF CULTURE ELEMENT DATA

BY

HAROLD E. DRIVER

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY, CALIFORNIA
1938

CULTURE ELEMENT DISTRIBUTIONS: VIII
THE RELIABILITY OF CULTURE ELEMENT DATA

BY
HAROLD E. DRIVER

ANTHROPOLOGICAL RECORDS
Vol. 1, No. 4

ANTHROPOLOGICAL RECORDS
EDITORS: R. H. LOWIE, A. L. KROEBER, R. L. OLSON
Volume 1, No. 4, pp. 205–220
Transmitted December 8, 1936
Issued April 20, 1938
Price, 25 cents

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY, CALIFORNIA

CAMBRIDGE UNIVERSITY PRESS
LONDON, ENGLAND

The University of California publications dealing with anthropological subjects are now issued in two series.

The series in American Archaeology and Ethnology, which was established in 1903, continues unchanged in format, but is restricted to papers in which the interpretative element outweighs the factual or which otherwise are of general interest.

The new series, known as Anthropological Records, is issued in photolithography in a larger size. It consists of monographs which are documentary, of record nature, or devoted to the presentation primarily of new data.

MANUFACTURED IN THE UNITED STATES OF AMERICA

CONTENTS

TABLES

CULTURE ELEMENT DISTRIBUTIONS: VIII
THE RELIABILITY OF CULTURE ELEMENT DATA

BY
HAROLD E. DRIVER

INTRODUCTION

Previous studies in reliability of ethnographical data are almost nonexistent. Many ethnographers have made brief remarks here and there about the reliability of a certain section or statement of a report but almost none has seriously coped with the problem. Informants have been characterized by catchwords such as good, poor, reliable, unreliable, and ethnographers as good observers, prejudiced observers, and the like.

The problem of the reliability of informants' statements raises at least three questions: How consistent is the same informant in response to the same questions on two or more occasions; what is the amount of agreement between two or more informants; and how near do they come to the actual facts? Material for answers to the first two queries is readily obtainable in the field, but that for the third is limited to live cultures or those for which we have exceptionally good historical documents or previous ethnographies.

So far as I know, the best study, to date, in reliability of native informants is that made by DuBois and Demetracopoulou[1] on the mythology of the Wintu Indians of California. They recorded numerous versions of the same myths in English, in the native language, through interpreters, from the same informant on two occasions a year apart, from informants of varying ages, and so forth. Their conclusions were roughly these: (1) Longer versions of the same myth differed from shorter versions mainly in the inclusion of songs and irrelevant ethnographnic detail and in repetition of the same elements rather than by the addition of definitely new elements. (2) "The presence of a native audience and confidence in the recorder frequently stimulate the narrator to tell longer tales." (3) Given the same careful field technique, English versions have essentially the same form and content as versions recorded in the native language. (4) "An able interpreter does not appreciably affect the content of a tale and may only slightly affect the form." (5) A lapse of time does not affect a story so much as the personality of the informant and circumstances attendant to the telling. Versions told a year apart differ no more than those told a few days apart. (6) The younger generation know far less mythology than their elders, but individual differences within a generation are greater than the differences of the norms of the two generations. (7) There were few local differences within Wintu territory.

Although these conclusions are indeed an important contribution to our knowledge of factors determining reliability of report, I believe they would have been sharpened by the use of a more definite analytical technique involving the splitting of the various versions of the same myth into elements or incidents which could then be compared numerically. Sheer length of the tale, however, is often given by some such phrase as "one-fourth longer."

The data of the present study were gathered much more rapidly and in a far less controlled manner and from fewer informants per tribe than were the Wintu myths. Nevertheless, they cover a much greater range of native culture and make possible the comparison of the reliability of various topics such as subsistence, games, shamanism, and the like. They consist mainly of responses to specific questions, whereas the Wintu myths are, of course, volunteer testimony. For these reasons, plus the difference in the nature of the material itself and the lack of general quantitative expression of the results of the Wintu study, no specific comparisons of the two will be made.

The data used in this study are mainly from culture-element lists obtained by Barnett[2] and myself.[3] Barnett worked two informants from Galice Creek in Oregon, and I two each from the Yurok, Hupa, and Karok tribes of northwestern California. Drucker filled in a column in Barnett's element list from his Tolowa data obtained previously by the usual field methods, to which was matched a list gathered by me from a single Tolowa informant. I also gleaned a few more items from Drucker's manuscript where they could be equated to elements in my list which had not appeared in Barnett's. Besides these, I read Kroeber's[4] account of the Yurok and those of Goddard[5] and Curtis[6] on the Hupa and entered the material (+ or -) into the element list used by me for the same tribes.

All the information employed in this study has been reduced to presence (+) and absence (-) form. Quite obviously this is a simplification of the true facts. In some contexts "+" means occurrence among a majority of the population, as in patrilocal residence. Matrilocal residence was

[1] 1932. For complete citation see References Cited.

[2] See Bibliography.

[3] CED:X--Northwest California, AR 1. (In press.)

[4] 1925. [5] 1903. [6] 1924.

practiced to some degree by all tribes in connection with "half marriage," the frequency among the Yurok being 23 per cent of 413 cases.[7] Nevertheless it was recorded "-" in the section on postnuptial residence. In other contexts "+" means present in any frequency at all, in some instances no doubt as low as one or two per cent. In still other sections, an element by its very nature would be restricted to some special group, such as sucking doctors. Unfortunately this factor of frequency has never been thoroughly discussed by any of us who have been gathering element-list material. So far no standardization has been attempted by me throughout my entire element list, but I did try to control the factor for each element individually in the field. In general, most elements present at all were recorded "+" by me. It was only, or at least mainly, where the item implied a majority, as in postnuptial residence, that the less frequent alternative was recorded "-." I do not hesitate to admit that this simplification is a short-cut device without other merit. The majority of information in monographs, however, is subject to the same limitations, which often are forced on the ethnographer by the limited knowledge of informants.

Concerning absences, few informants consistently deny the presence of elements unknown to them. They often say, "I never saw it," "I never heard of it," or "My father never told me about it." In such a situation, I usually recorded minus. Wherever I felt certain that the informant was mistaken I entered a query or wrote a note expressing my doubt. Where either the informant or myself seemed to be in doubt, I enclosed the entry in parentheses. Parenthetical entries have been included in the statistics to follow, because if they are inferior items, it is desirable to know how they behave. Not infrequently an informant would recognize my description or illustration of an element as belonging to a neighboring tribe and deny it for his own without hesitation. Other references to field technique will be made below from time to time.

While the concrete discussion to follow is limited to four tribes, the element universe extends from the Tillamook in Oregon to the Kato in California. In other words, any element present in any of the twenty-three tribes[8] in this area was included in at least some statistics below. Had a wider areal universe been chosen, it would have increased the number of common absences (--) among the four tribes and raised correlation coefficients. A narrower universe would have lowered such values.

In most of the statistical treatment to follow I have compared informants with informants. The culture elements are the units counted. Were there more duplicated bodies of data available, the problem might have been approached wholly from the other side, by determining how many pairs of informants agree in their responses to each element, and percentaging these values to arrive at a measure of the reliability of each element. This has been done for a part of the data but is inconclusive because of the small quantity of duplicated tribal inventories.

To determine the reliability of any body of source material it is necessary to use some statistical measure. It is desirable to employ the same measure to be used later on to correlate the material so that the results will have bearing on the intertribal or intertrait correlations. Thus if one expects to apply Q_β to a number of tribal inventories to obtain groupings of tribes (culture "areas"), he should use this same formula to determine the amount of correlation or association between the responses of two informants of the same tribe. If he finds, e.g., that a number of pairs of informants show intratribal correlations of about .90, this means that intertribal correlations of the same magnitude, if any, are likely to have true values of 1.00. If it can be shown that intratribal differences, which are errors on the part of at least one of the informants, are randomly distributed throughout a list of elements, then all obtained correlations will automatically be lower than their true values and can then be raised to their true values by a correction device.[9]

I have chosen tetrachoric r (r_t),[10] computed graphically from Thurstone's diagrams,[11] as a measure of the correlation between the responses of pairs of informants.

Most of the fourfold distributions are nearly symmetrical[12] and any other measure would give highly parallel results. Under conditions of perfect symmetry $Q_\beta = r_t$. However, where an appreciable amount of asymmetry is present, I believe Q_β or the r included in it are more satisfactory than r_t.

Table 3 (see at end of this study) gives the standard error of r_t. When the difference between two correlations is 2.5 times the standard error of the difference, $\sqrt{\sigma r_{t_1}^2 + \sigma r_{t_2}^2}$ the chances are approximately 99 to 1 that it is real, i.e., not owing to mere sampling error. In general, I believe it legitimate to assume that any number of elements is a sample of some very much larger totality. However, where the number of possibilities is definitely limited, as in fire-making methods or arrow releases, we cannot consider the 5 or so possibilities as being mere samples of a very large number of alternative methods. Differences in informants' testimony which occur in such universes are real if only five elements are

[7] Waterman and Kroeber, 1934.

[8] All tribes in Barnett's and my own distribution studies.

[9] Cf. Spearman, 1904, 1907.

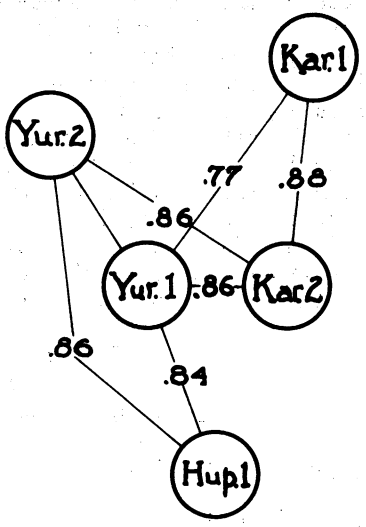
[10] Pearson, 1900.

[11] Chesire, Saffir, Thurstone, 1933.

[12] I.e., the percentage ratio of positive to negative responses is about 50:50.

involved. From a broader point of view, however, such elements can be considered samples of primitive technology, and ultimately as samples of the total number of kinds of human cultural behavior. Sampling concepts seem to apply to these broad categories, not to very small ones. In some of the arbitrary sections into which I have divided the data of tables 3 and 4, the elements constitute a large part of the total number of comparable variants in the area: e.g., houses or games. Other sections, such as ceremonies and shamanism, perhaps represent a smaller fraction of the total number of possible comparable details. The reality of numerical differences throughout the sections to follow will be subjectively evaluated from time to time, but in general I shall follow the usual sampling theory.

While the source material apparently offers duplicated information on five tribes, pairs of informants from two of these, Karok and Yurok, belong to different local groups. The first problem is to decide whether there is any justification for assuming that the two Karok and the two Yurok informants respectively belong to the same cultures. Certain relationships are given in the figure.



Although the two Karok groups are only slightly more disparate than the two Yurok, the external relations show a definite difference. The correlation of Karok 1 to Yurok 1 is definitely lower than that of Karok 2 to Yurok 1. This means that all the differences between Karok 1 and Karok 2 are not randomly distributed with reference to Yurok 1, as errors would be, but that some of them behave systematically. Karok 1 is more aberrant culturally as well as geographically than Karok 2. In contrast, the external relations of the two Yurok groups to Karok 2 and Hupa 1 are nearly the same. The differences between the Yurok groups are randomly distributed with reference to two neighboring localities and thus behave like errors. Because of these facts I have eliminated the Karok from the rest of the study.

Informants from the Galice and Hupa tribes are certain to belong to culturally identical local groups. The settlements of each of these tribes all fell within a 10-mile length of stream. In contrast the Yurok and Karok occupied some 40 and 50 miles of river respectively. The informants from Galice creek are blood relatives, "cousins." Those for Hupa were unrelated but were from villages only 3 miles apart which joined together for certain ceremonies. For the Tolowa, Drucker used a number of informants, including the man who served as my only informant. Thus Drucker's material is perhaps less localized than my own, but this difference is slight because the Tolowa were a small group compared to the Yurok or Karok.

It should be remembered throughout that the three ethnographers, Drucker, Barnett, and myself, did not work under parallel field conditions. Drucker spent 70 or 80 days with a number of Tolowa informants gathering as full material as was available. Barnett began with a relatively short and generalized element list (less than 1000 elements) constructed chiefly by Kroeber for all of California, and built this up with new elements obtained from informants to about 2000. He spent a week each with his Galice informants. I began with a list of about 3000 traits, made some inquiry into all of these, added more while in the field, and spent only three or four days with each informant. Neither time nor method of inquiry were held constant by the three field investigators.

RELIABILITY OF TRIBAL TOTALITIES

The totals of tables 1 and 2 summarize the chief findings. Table 1 includes all information. The values for N differ from tribe to tribe because we happened to have more duplicated information on the Yurok, e.g., than on the Tolowa. The 1366 Tolowa elements are not all included in the 2337 for the Yurok, which means that the correlations throughout this table are based on different universes for the different tribes. These universes are not strictly random samples of a larger totality, hence the correlations in this table are not strictly comparable to one another. Table 2 gives the results from the 706 elements which were positively or negatively reported for all four tribes. These correlations, made from the same universe, are wholly comparable.

The correlation coefficients (r_t) are these:

Tribes	Table 1	Table 2
Galice	.92	.92
Tolowa	.87	.90
Yurok	.91	.97
Hupa	.97	.98
Average	.92	.94

The correlations based on table 2 are higher than those of table 1, but hardly significantly so except with respect to the Yurok. I can think of no explanation whatsoever for the difference in the two Yurok values. For the phenomenon as a whole, the apparent reason is that the shorter element list (table 2) represents elements of more general knowledge. These elements are what all ethnographers (Barnett, Drucker, and myself) thought worth including in the list of queries, and at the same time what informants would readily give answers to or sometimes volunteer. Such would expectably be more reliable.

In both tables the Hupa informants show a very high reliability. The second informant was hard of hearing and slower in response than the first. I suspect I tended to equate hesitant or uncertain responses to those of the first informant. I had positive assurance that both belonged to culturally identical local groups and should agree throughout. Furthermore, they were the first two informants I worked with in the area, and my knowledge of the culture and ability to control its elements was less than for the other tribes. I do not believe their true knowledge is sufficiently superior to other informants' to account for this difference.

Barnett's Galice creek material does not reveal this type of bias, at least in comparison with correlations of the other three tribes.

The relatively low reliability of the Tolowa is doubtless attributable to the fact that there were two ethnographers, Drucker and myself. Drucker worked without an element list, using the more usual method of inquiry where more information is volunteered. He may have erred occasionally in entering negatives, some of which were perhaps inferred from his general knowledge. However, having read his entire manuscript, I am convinced that the vast majority of differences are my errors. Furthermore, Drucker spent some 70 or 80 days with Tolowa informants, I 3 days.

Further evidences of the influence of the ethnographer on correlations are these:[13]

Sources	++	+-	-+	--	N	r_t
Drucker's Tolowa-Galice 2	367	97	68	352	884	.84
Driver's Tolowa-Galice 2	363	119	74	328	884	.80
Drucker's Tolowa-Yurok 2	382	83	146	273	884	.70
Driver's Tolowa-Yurok 2	411	70	114	289	884	.81

Our two Tolowa lists are compared with the Galice on the north and the Yurok to the south. I had worked with a Yurok before visiting the Tolowa. Drucker had not. His information on the Tolowa and that of Barnett on the Galice are essentially independent of each other although they discussed some of the Galice material at the time Barnett was gathering it. My Tolowa information correlates .09 higher than Drucker's with the Yurok, and .04 lower with the Galice. It seems obvious to me that I projected previous knowledge, mainly Yurok, into my Tolowa data. The fact that my informant came from the southern part of the Tolowa territory nearest the Yurok, and that there were a number of admitted Yurok influences present in small frequencies, may have caused some of the foregoing differences.

Further evidence of the rôle of the ethnographer is given by these figures on four Hupa sources:

Sources	++	+-	-+	--	N	r_t
Goddard-Curtis	295	6	5	27	333	.97
Goddard-Driver 1	274	25	5	29	333	.90
Goddard-Driver 2	274	25	5	29	333	.90
Curtis-Driver 1	270	30	8	25	333	.85
Curtis-Driver 2	270	30	8	25	333	.85
Driver 1-Driver 2	274	4	4	51	333	.99

The values .85 and .90 are not significantly different because the standard error of the difference is about .04. The value .97 is just barely significantly higher than those of .90. Thus the significant facts are represented by the two higher correlations on the one hand and the four lower ones on the other. The very high agreement between my own informants has been interpreted above. Concerning the relationship between Goddard's and Curtis' accounts, I am convinced that Curtis drew from Goddard when he wrote his report. The following tabulation of positive elements (+) mentioned in their works proves this.

		Goddard: Mentioned	Goddard: Not mentioned	Totals
Curtis	Mentioned	295	88	383
	Not mentioned	228	478	706
	Totals	523	566	1089

The total, 1089, is the number of positive entries obtained by me from the Hupa. Negatives were disregarded because of their infrequency in the published sources. The correlation of this fourfold table is $r_t = .64$. Had Goddard and Curtis randomly selected from 1089 positive elements, the correlation would have been $r_t = .00$. Perhaps some of this agreement in selection of material is due to the kind of information volunteered by informants, but I doubt if such a high agreement would have come about in this way. The actual terms and sentence structure of Curtis conform so closely to Goddard that the independence of the reports could be doubted on these grounds alone. Furthermore, Curtis' chapter on

[13] The raw frequencies have been included in the small tables in the text in order to give a sample of their distribution and the amount of asymmetry present.

the Hupa is longer than any other chapter in his work on tribes in the general area. At the time he wrote, Goddard's report was practically the only published literature available. Curtis does not mention names of his informants. These facts all point to one conclusion.

Within the limits of the 333 examples cited above, what superficially appeared to be four sources thus turn out to be two. The low values, .85 and .90, which represent the correlations between these two sources are of about the same magnitude as those between other more or less independent sources (totals of tables 1 and 2), and do not require any special explanation.

For fear that the 333 traits above might have been a somewhat selected rather than representative sample of the larger universe, I made another count of all elements reported by Goddard and myself. These total 613:

Sources	++	+-	-+	--	N	r_t
Goddard-Driver 1	467	56	11	79	613	.90
Goddard-Driver 2	470	53	8	82	613	.93
Driver 1-Driver 2	468	9	9	127	613	.99

They show no significant differences from the 333 discussed above.

Further relationships between element list and monographic sources are found by comparing Kroeber's (1925) account of the Yurok with the two Yurok lists obtained by me.

Sources	++	+-	-+	--	N	r_t
Kroeber-Driver 1	549	43	33	206	831	.95
Kroeber-Driver 2	536	58	29	208	831	.94
Driver 1-Driver 2	526	56	37	212	831	.93

These correlations do not differ significantly from one another. Therefore my second Yurok list was probably not directly influenced by my first. Incidentally, I interviewed the Tolowa informant between the first and second Yurok. The lapse of a few days may have contributed to the independence of results.

If more informants had been used it would have been possible to determine the influence of sex, age, native occupation (e.g., shaman), etc., on report. Had more field controls been employed, the reliability and total quantity of volunteer versus questioned testimony, time held constant, could have been determined. This is difficult to accomplish from published literature owing to the dearth of negative evidence and specific field controls.

Within the limits of the variation among ethnographers and informants in this study, the personal equation of the ethnographer seems to be as important a determinant of reliability as that of the native informant.

RELIABILITY WITHIN SPECIAL TOPICS

The topical arrangement of tables 1 and 2 is the same as that used in the field except that two or three sections have been combined occasionally to raise the total number of elements (N) to a figure large enough to minimize sampling error. Therefore the classification is not influenced by any desire to "prove" preconceived theories. Although the four reliability coefficients within a given class sometimes show appreciable differences, the following arrangement of averages in rank order provides a summary for the four tribes as a whole.

Correlation (r_t) Averages of Various Topics

Topics	Table 1	Table 2
Ceremonies	.98	
Marriage	.96	.99
Houses	.96	
Games	.95	.98
Money, tobacco, musical instruments	.94	
Weaving	.91	
Counting, astronomy	.91	
Death	.89	
Body and dress	.88	.94
Birth, menstruation	.87	.88
Navigation, technology, weapons	.86	.89
Social stratification, war	.85	
Shamanism	.85	.85
Subsistence	.80	.86

The blank spaces in the table 2 column are due to the fact that no correlations were computed when N was under 50.

While the range of this series of correlation averages is .18, there are no definite breaks within it. The variation in reliability of these 14 topics is not greater than that of the tribal totals discussed above, except for subsistence (.80). On the whole, material culture is no more reliable than social or religious culture. The general belief that material culture is more reliable than other sections no doubt comes from the fact that it is preserved in museums and sometimes can be photographed in the field. A novice attempting to describe living material culture is likely to accomplish more than he would with social organization. When, however, almost the entire culture is gone, and almost all information has to be salvaged from the memories of native informants, social and religious concepts can be obtained with about the same reliability as material ones.

The high reliability of ceremonies is explained partly by the fact that they are still given by some of the tribes. They are also completely exoteric except for a few magical formulae recited by a priest. Furthermore, the routine of dancing and singing is repeated in identical form for

eight or ten days. In recent years these ceremonies have been biennial for the Yurok and Hupa instead of annual, but even so the average informant has seen them twenty or thirty times.

The high average reliability coefficient for marriage is attributable partly to the fact that parent-in-law child-in-law avoidances have been included in this section. These avoidances, totaling 36 traits, are entirely absent for all four tribes. Thus the common absence category of each relationship is increased by 36. This raises correlations. The elimination of avoidances would lower the average correlation to about .90.

Concerning houses, a number of Yurok and Hupa structures are still standing. I filled in some of the entries from direct observation. Where the two informants were of the same local group, as at Hupa, I eliminated my own direct observations from the statistics. Under such a condition of complete preservation, informants' descriptions would also be highly accurate.

Concerning subsistence, the relatively low reliability in table 1 can be explained by at least two factors. First, minor differences in geographical environment and ultra-localization by the informant undoubtedly introduce a number of real differences which are not errors on the part of either of two informants. This would apply mainly to the Yurok, although I eliminated from the count a few elements which obviously showed true local differences. Furthermore, I always began my inquiry with subsistence and proceeded in the order of tables 1 and 2. Many informants, anxious to please, volunteered more information in the first few hours of the interview than during the rest of it. I tended to introduce a good part of this into the element list and expanded the subsistence section to the point of diminishing returns. The significantly higher reliability from the smaller sample of table 2 corroborates this interpretation. It is obvious that the quantity of detail concerning any topic can be increased to the point where an average informant will make a high percentage of error. A small quantity of information on a given topic is also unsatisfactory because it is likely to be an insufficient sample of the informant's total knowledge and of the total number of relevant facts. The desideratum lies somewhere between. However, the relatively low correlation of two Tolowa informants is scarcely explainable on environmental grounds because the area was small, and neither does it fit very closely the over-expansion theory because the total number of subsistence elements is only 172 compared to 322 for the Yurok. The difference must therefore also have been caused by the particular knowledge of the ethnographers and informants involved.

Such interpretations could be extended to others of these topics but would become increasingly subjective and have little bearing on other areas or cultures.

Several other divisions of the total body of data have been made. One of these is material objects versus the rest of the list. Here a sharper distinction than any cited above between material objects and behavior or belief has been made. While a topic like subsistence includes many material objects, perhaps half of its elements refer to some kind of behavior or belief connected with obtaining or preparing food. It is therefore not composed entirely of material objects. The present classification attempts to include only such elements as could be photographed or observed in museum or field without any information whatsoever regarding manufacture or use. Examples: all weapons are included but whether they are used in war or hunting is ruled out; hair coiffure and tattooing are included, but the washing or greasing of hair and method of tattooing are omitted; gaming objects are included, but not the rules of play; structural features of houses are accepted but not the fact that men habitually slept in the sweat house. Tribal correlations of such elements follow. They are compared with the totals of table 1.

Pairs of informants	Material objects						Total elements
	++	+-	-+	--	N	r_t	r_t
Galice	193	28	33	217	471	.92	.92
Tolowa	257	50	35	169	511	.85	.87
Yurok	412	63	46	279	800	.91	.91
Hupa	422	20	13	280	735	.99	.97

There are no significant differences between these two sets of correlations. These facts, combined with the ranked list of topics above, are proof that material culture elements or material objects are no more reliable than other elements.

A number of illustrations were shown by me to informants in the field. This was done for both Hupa informants, both Yurok, and the one Tolowa. So far as I know, Drucker and Barnett used no illustrations. The question arises: are the responses obtained from pictures more reliable than those obtained from verbal questions alone? These are some of the relevant facts:

Informants	Illustrated elements					
	++	+-	-+	--	N	r_t
Yurok	217	18	6	54	295	.95
Hupa	225	10	4	57	296	.98

The Yurok correlation here is barely significantly larger than the Yurok figure (.91) for material objects and the total list given above. The Hupa remains the same. The fact that illustrated elements show a slightly higher correlation for the Yurok is no proof that the pictures caused this difference. A proof of the efficacy of pictures demands a control group of informants on whom no pictures were used. This control group would have to be identical in knowledge, command

of language, etc., with the group shown the illustrations, so that whatever difference occurred could be attributed definitely to the pictures. These conditions are far from satisfied by these data. Nevertheless, a comparison of the reliability of the 141 illustrated elements which are reported for all four tribes with the reliability of the total body of data and material objects seems worth while.

Informants	++	+-	-+	--	N	r_t
Galice	62	5	8	66	141	.95
Tolowa	74	19	3	45	141	.91
Yurok	108	7	3	23	141	.95
Hupa	111	4	2	24	141	.98

The fact that the Galice correlations are also higher for pictured elements when no pictures were used eliminates the significance of the Yurok value. From these scraps of evidence there is no proof that illustrations improve reliability. Personally, I believe they do help. No one knows how much time either Barnett or myself spent in obtaining information on these 141 items. This would have to be held constant before the efficacy of pictures becomes determinant.

It occurred to me that more or less generic traits might show a higher reliability than more specific traits, the theory being that one forgets more details than general facts. The definition of generic versus specific is, of course, a relative one, but in order to avoid selection in favor of a preconceived theory I have almost mechanically followed indentations in my original field manuscript. For example:

Wooden chest
  Cylindrical
  Rectanguloid
  In two pieces, about equal size
  Small opening and lid
  Carved decoration
  For feathers and valuables

The indented elements all refer to the wooden chest. A generic element, then, is one which has two or more specific subvariants; a specific element, one of two or more subvariants of a generic element.

Facts concerning the total number of generic and specific elements derived from the entire list (the universes of table 1) are:

	Generic elements						Total elements
Informants	++	+-	-+	--	N	r_t	r_t
Galice	128	9	11	46	194	.93	.92
Tolowa	147	21	12	33	213	.80	.87
Yurok	209	16	10	58	293	.94	.91
Hupa	176	6	2	65	249	.99	.97
Average						.92	.92

	Specific elements					
Informants	++	+-	-+	--	N	r_t
Galice	255	31	63	367	716	.91
Tolowa	294	48	50	281	673	.90
Yurok	537	103	62	554	1256	.92
Hupa	503	51	30	517	1101	.97
Average						.93

The only statistically significant differences in these three sets of correlations are those of the Tolowa. I have no special explanation for these differences. From the facts as a whole, neither generic nor specific elements show any significant differences in reliability from the entire list.

It occurred to me further that those elements which were widely distributed in an area might be more reliable than those of limited distribution. Presumably, the more widely distributed elements would occur in higher frequencies in individual tribes and hence play a more prominent rôle in the culture. To make the contrast sharp, I chose as widely distributed elements only those which were present in 11 or more of the 15[14] tribes in the area. Elements of limited distribution are defined as those known to be absent in 11 or more of the 15 tribes but present in at least one of the four duplicate tribes. These are the findings:

	Wide distribution							Total elements	
Informants	++	+-	-+	--	N	W	r_t	W	r_t
Galice	188	9	23	21	241	.87	.76	.87	.92
Tolowa	263	23	5	13	304	.91	.72	.84	.87
Yurok	401	19	14	1	435	.92	.03	.86	.91
Hupa	419	12	2	4	437	.97	.65	.91	.97
Average						.92	.54	.87	.92

	Limited distribution						
Informants	++	+-	-+	--	N	W	r_t
Galice	9	4	5	54	72	.88	.82
Tolowa	31	9	12	25	77	.73	.65
Yurok	21	28	36	64	149	.57	.10
Hupa	20	8	3	120	151	.93	.94
Average						.78	.63

Because of the highly asymmetrical nature of the fourfold distributions involved, I have computed the percentage of agreement, W, as well as the correlation coefficient.[15] Widely distributed elements show a slightly higher percentage of agreement between pairs of informants than the total list (table 1). Elements of limited distribution show a definitely lower percentage of

[14] The areal universe is that of table 4. The Karok and Sinkyone are counted as 2 tribes each. The rest of Barnett's material was in press, and not available to me at the time this count was made.

[15] W = (a+d)/N where a is the number of elements for which both informants gave positive responses, d the number for which both gave negative responses, N the total number of elements.

agreement than the total list. Correlations give a very different picture. The total list is highest, next elements of limited distribution, and finally those of wide distribution. Because the type of selection automatically produces asymmetrical distributions, the meaning of the correlations becomes problematical. Therefore I am accepting the results from W as being the more meaningful. Thus widely distributed elements are more reliable than those of limited distribution.

To summarize the findings of this section: (1) While there are definite differences in the reliability of groups of elements arranged topically, the range of variation of the reliability coefficients is no greater than that of those of the four tribal totalities, and there are no definite breaks in the series. (2) Material culture or material objects show the same degree of reliability as the entire element list. (3) There is no proof that illustrations improve reliability. (4) Neither generic nor specific elements show any greater or lesser reliability than the list as a whole. (5) Widely distributed elements within a given area are apparently more reliable than narrowly distributed elements.

RELIABILITY OF INDIVIDUAL ELEMENTS

The reliability of an individual element can be measured in terms of the number of pairs of informants who agree in their responses to it. In the present example, when all four pairs of informants show intra-agreement the element is completely reliable. Table 4 is a compilation of the presumably least reliable elements, those for which two, three, or four pairs of informants disagreed. The element numbers are those of my field study.[16] The universe is that of table 2, N=706. A few additional elements have been included to give supplementary information here and there.

It is significant that only a single element shows four differences, and only nine show three differences. The distribution of these differences among the 706 elements is the following:

Element frequency	Number of pairs of informants with intradifferences				
	0	1	2	3	4
Actual frequency	477	175	44	9	1
Chance frequency	455	211	36	3	0

The chance frequencies are those which would come about if the 294 differences were randomly scattered among the 4(706) = 2824 duplicated entries.

I am indebted to Dr. J. M. Thompson for showing me the following method of deriving the chance frequencies. Let a be the probability of a difference occurring between any pair of informants for any element. Then:

[16] Driver, MS in press.

a^4 = the probability that all four pairs of informants will intradiffer on the same element.

$4\,a^3(1-a)$ = the probability that exactly three pairs of informants will intradiffer on the same element.

$6\,a^2(1-a)^2$ = the probability that exactly two pairs of informants will intradiffer on the same element.

$4\,a(1-a)^3$ = the probability that exactly one pair of informants will intradiffer on the same element.

$(1-a)^4$ = the probability that no pair of informants will intradiffer on the same element.

Since N = 706 elements, the total number of duplicated responses is 4(706) = 2824. The total number of differences, by actual count, is 294. Therefore:

$$a = \frac{294}{2824} = .104$$

a^4 = .000, which, multiplied by 706 = 0
$4\,a^3(1-a)$ = .004, which, multiplied by 706 = 3
$6\,a^2(1-a)^2$ = .052, which, multiplied by 706 = 36
$4\,a\,(1-a)^3$ = .299, which, multiplied by 706 = 211
$(1-a)^4$ = .645, which, multiplied by 706 = 455
1.000 705

The frequencies of elements which show two, three, or four differences are so near chance, that we have practically no proof that they are poor elements. The accumulation of differences is therefore mainly due to unknown factors whose cumulative effect produces distributions similar to those of coins or dice. Such factors might be: error in recording on the part of the ethnographer; verbal error in response on the part of the informant; or true misunderstanding on the part of either party. The fact that the questions are never given in exactly the same words by the ethnographer on two occasions means that the stimulus is not fully controlled. Such control will probably never be achieved because the variety of cultures and languages involved is too great and the time too short to realize this ideal. A difference in the rôle played by an element in a culture may make an ethnographer's query meaningless to certain informants. In my experience with this type of interview I have done at least half of the talking. Each element must be described in some context before an informant will get the idea, and the ethnographer is forced to decide whether or not his description to the informant is adequate when the informant continues to respond negatively. Under such conditions of work, random error would be expectable.

On the assumption that at least some of the ten elements for which three of four intratribal differences occurred are inferior items, I shall give my impressions regarding some of the possible

causes of these differences. Numbers refer to elements in table 4.

5. Informants were doubtless confused between the construction of a definite brush fence and the piling of a little brush around a natural trail or runway. This element also appears to be more typical of southern Athabascans than the Klamath and Smith river region where systematic fishing was far more profitable than small game hunting.

27-28. No doubt a secondary adaptation of a dip net, and only one of several ways to obtain woodpeckers. They were also shot and snared.

88. Individual differences within a local group are certain to have occurred. Some persons ate such animals, others did not.

423. Food was nowhere habitually sold. Within a family or small village it was communal. Without, it was bartered for other natural products or articles of value, occasionally dentalia. The caption lacks specificity.

1550. A child's toy, hence of small importance. It also seems to be more typical of central California.

1648. This was no doubt of uncommon occurrence and resorted to only when the girl's family was in desperate need of money.

1920. There seems to have been appreciable variation in the age at which children were named. There may also have been a confusion between a mere nickname and one formally bestowed.

1961. In many instances the girl was supposed to pick up a few small sticks of wood on her way home from her daily bath in the stream. She went outdoors at no other time. This was doubtless interpreted as work by some informants, but not by others.

2282. Red is only one of several colors used. Face painting was never symbolic in northwestern California, and the color used perhaps subject to local availability or individual preference.

In general, these "worst" elements appear to be unimportant or of infrequent occurrence in northwestern California culture. In order to prove that this is the cause of their unreliability one would have to determine the frequencies of "unimportant" or "infrequent" items throughout the entire list and see if intratribal differences accumulated in greater proportions in these items than in the rest of the list as a whole. This would be a difficult and subjective procedure and I shall not attempt it.

Unknown causes (chance) seem to be more potent determiners of the reliability of individual items than known causes.

PRACTICAL RECOMMENDATIONS

While these findings answer a few questions concerning the reliability of culture-element data, we still know relatively little about the reliability of individual items, which is the one thing we want most to know if we are going to improve our prefield element lists. If even 25 per cent of the tribal element inventories already collected were duplicated by a second ethnographer and second informant, we would have material enough to come to definite conclusions. With fifty or so duplicated lists randomly distributed over the area of our present activity, we could empirically determine the reliability of each element with some little assurance. I see no point in undertaking intensive studies of the DuBois-Demetracopoulou type to solve the element list problem. These are expensive, good informants would drop dead around us while we were working with a single tribe, and the results would not be directly applicable to the culture-element survey.

If, from a large body of duplicated data, we still found that errors tended to be randomly distributed, we would have ample justification for correcting all obtained correlations upward to their true levels. Only local (topically as well as areally local) correlations would then differ to any important degree from their true values.

SUMMARY AND CONCLUSIONS

Compared with data of other social sciences, the reliability of culture-element material is fairly high. Reliability coefficients, r_t, of the entire body of data examined range from .87 to .97. Percentages of agreement vary from .84 to .91. The low values are likely to be nearer the true reliability because higher ones are from data collected by a single ethnographer who may have possessed a bias of some kind.

Within the limits of personal variation among the present ethnographers and informants, the personal equation of the ethnographer seems to be as important as that of the informant.

Although certain topics are definitely more reliable than others, the range of variation is little greater than that of total tribal inventories. No one major division of culture, material, social, or religious, shows any higher reliability than any other.

Material objects show the same reliability as the entire body of data.

Illustrations apparently do not increase reliability.

Generic or specific elements show no significant differences in reliability from each other or from the data as a whole.

Widely distributed elements are apparently more reliable than the unselected elements in the list; narrowly distributed elements less reliable.

Individual elements show few demonstrable differences in reliability.

There is a need for more duplicated source material if we are to learn more about the kind and cause of differences between informants and ethnographers.

TABLE 1

Intratribal Correlations from All Data

(viz., Galice Inf't 1 with Galice Inf't 2, etc.)

	Correlations (r_t)					Number of elements (N)			
	Galice	Tolowa	Yurok	Hupa	Average	Galice	Tolowa	Yurok	Hupa
Subsistence	.87	.63	.74	.95	.80	162	172	322	315
Houses	.96	.91	.98	1.00	.96	177	192	226	88
Navigation, technology, weapons	.74	.82	.89	.98	.86	126	131	240	222
Body and dress	.96	.67	.89	.99	.88	93	94	208	196
Weaving	.77		.96	1.00	.91	57	46	105	105
Money, tobacco, musical instruments	.98	.92	.88	.98	.94	73	64	84	84
Games	1.00	.96	.96	.89	.95	89	85	155	120
Counting, astronomy			.84	.98	.91	19	15	76	56
Marriage	1.00	.97	.89	.98	.96	91	94	142	138
Birth, menstruation	.84	.80	.92	.92	.87	134	121	203	166
Death	.82	.96	.83	.93	.89	79	85	142	123
Social stratification, war	.92	.56	.95	.96	.85	104	56	126	100
Shamanism	.88	.63	.96	.93	.85	132	94	171	146
Ceremonies	.97	1.00	.96	*	.98	124	117	137	*
Total	.92	.87	.91	.97	.92	1460	1366	2337	1859

*I witnessed the White Deerskin Dance and obtained information about other ceremonies from only one informant.

TABLE 2

Intratribal Correlations from 706 Elements

	Correlations (r_t)					N
	Galice	Tolowa	Yurok	Hupa	Average	
Subsistence	.86	.75	.88	.94	.86	93
Houses						18
Navigation, technology, weapons	.76	.85	.93	1.00	.89	82
Body and dress	.98	.86	.92	1.00	.94	56
Weaving						35
Money, tobacco, musical instruments						49
Games	1.00	.98	1.00	.95	.98	64
Counting, astronomy						6
Marriage	1.00	1.00	.97	1.00	.99	69
Birth, menstruation	.84	.86	.97	.85	.88	71
Death						40
Social stratification, war						31
Shamanism	1.00	.44	.99	.97	.85	50
Ceremonies						42
Total	.92	.90	.97	.98	.94	706

TABLE 3

Standard Errors of Tetrachoric R

(Standard errors of r_t, o r_t, when asymmetry* = .00 for both variables. These increase at an increasing rate as asymmetry of either variable increases. They are increased 50 per cent when asymmetry = ± .70 for both variables, and are doubled when asymmetry = ± .82 for both variables. Compiled from Pearson, 1913.)

	Correlation Coefficient, r_t												
N	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95	.99
50	.197	.190	.181	.172	.162	.151	.137	.124	.107	.089	.068	.042	.015
75	.160	.154	.148	.140	.132	.123	.112	.101	.088	.073	.055	.034	.011
100	.139	.134	.128	.122	.115	.107	.097	.087	.076	.063	.048	.030	.009
150	.113	.109	.105	.099	.093	.067	.079	.071	.062	.051	.039	.024	.008
200	.098	.095	.091	.086	.081	.075	.069	.062	.054	.045	.034	.021	.007
300	.081	.078	.074	.071	.066	.062	.057	.051	.044	.037	.028	.017	.005
400	.070	.067	.065	.061	.058	.054	.049	.044	.038	.032	.024	.015	.005
500	.052	.060	.057	.054	.051	.048	.043	.039	.034	.028	.021	.013	.004
700	.052	.050	.048	.046	.043	.040	.037	.033	.029	.024	.018	.011	.003
1000	.044	.043	.041	.039	.037	.034	.031	.028	.024	.020	.015	.009	.003
1500	.036	.035	.033	.032	.030	.028	.025	.023	.020	.016	.013	.008	.002
2000	.031	.030	.029	.027	.025	.024	.022	.019	.017	.014	.011	.007	.002

*Asymmetry is defined as $\frac{2\Sigma+}{N}$ - 1.00 where plus refers to the positive responses.

TABLE 4

Distributions of Least Reliable Elements

(+, present; -, absent; Gal 1, 2, Barnett's first and second Galice informants; Tol Di, Driver's Tolowa; Tol Du, Drucker's Tolowa; Yur 1, 2, Driver's first and second Yurok informants; Hup 1, 2, Driver's first and second Hupa informants; numerals to left of columns of + and - entries = number of pairs informants who differed.)

		Gal 1	Gal 2	Tol Di	Tol Du	Yur 1	Yur 2	Hup 1	Hup 2
Subsistence									
Hunting									
1. Driving into fence with nooses in gaps		+	+	+	+	+	-	+	+
5. Rabbits	3	+	-	+	-	-	-	-	+
6. Quail	2	+	+	-	+	-	-	-	+
8. Driving with fire		-	+	-	-	-	-	+	+
11. Small game	2	-	+	-	-	-	-	-	+
13. Driving into water	2	+	+	+	+	+	-	-	+
27. Nets, bag type	3	+	-	+	-	+	+	-	+
28. Woodpeckers	3	+	-	+	-	+	+	-	+
39. Deadfalls		+	+	+	+	+	+	+	-
44. For large game	2	-	+	+	+	-	+	-	-
45. For small game	2	-	+	+	+	+	+	+	-
Animals Eaten									
88. Ursus horribilis eaten	3	+	-	+	-	-	+	+	+
89. Felis cougar eaten	2	+	+	-	-	-	+	+	-
Fishing									
175. Gill net	2	-	+	+	+	-	+	+	+
190. Crab-claw rattle on net	2	-	+	+	+	-	+	-	-
Various									
423. Food sold	4	-	+	+	-	+	-	+	-
Houses									
588. Notched plank or log ladder		-	+	+	+	+	+	+	+
589. Type b house	2	-	+	+	+	+	-	-	-
Technology									
747. Wood meat platter	2	+	-	+	+	+	+	+	-
Weapons									
879. Wooden arrow straightener			+	+		+	+	+	+
880. Perforated				+		+	+	+	+
881. Forked stick	2	-	+	+	-	+	+	+	+
895. Quiver carried at side under arm	2	-	-	+	-	-	+	+	+
896. Quiver carried on back				+		+	+	-	-
907-8. Slings used	2	+	+	+	-	+	-	-	-
906. For hunting, by men	2	-	+	-	-	+	-	-	-
909. For war		-	+	-	-	-	-	-	-
910. As boy's toy only		+		+	-	-	-	-	-
922. Elkhide helmet	2	-	+	+	+	+	+	-	+

TABLE 4 -- Continued

		Gal 1	Gal 2	Tol Di	Tol Du	Yur 1	Yur 2	Hup 1	Hup 2
Body and Dress									
961. Soaproot brush for hair	2	-	-	+	-	-	-	-	+
986. Cap of fur for men	2	+	+	+	-	+	-	-	-
994. Hide shirt	2	+	+	+	-	-	+	-	-
995. Buckskin	2	+	+	+	-	-	+	-	-
1001. Breechclout (between legs)	2	+	+	+	-	-	+	-	-
1002. Buckskin	2	+	+	+	-	-	+	-	-
1037. Leggings	2	+	+	+	-	-	+	+	+
1038. Buckskin	2	+	+	+	-	-	+	+	+
Weaving									
1163. Deep sifting of winnowing basket, pointed bottom	2	-	-	+	-	+	-	+	+
1241. Net mesh spacer of wood	2	+	+	-	+	+	+	+	-
1242. Net mesh spacer of bone or horn				+	+	+	+	+	+
Musical Instruments									
1311. Hide drum	2	+	+	+	-	+	-	-	-
1340-1. Bull-roarer	2	+	-	-	+	-	-	-	-
1345-6. Toy	2	+	-	-	+	-	-	-	-
Games									
Hand, Grass, Many Stick Game									
1452. Two bones or sticks per player	2	-	+	-	-	-	-	+	-
1454. Of hollow bone	2	-	+	-	-	-	-	+	-
1464. Hide in grass in hand	2	-	+	-	-	-	-	+	-
1545-6. Jacks	2	-	+	-	+	-	-	-	-
1547. With stones	2	-	+	-	+	-	-	-	-
Tops									
1550. Acorn	3	-	-	+	-	+	-	-	+
Marriage									
1647. Child betrothal before puberty		+	-	+	+	+	+	-	-
1648. With payment by groom's side	3	+	-	-	+	-	+	-	-
Birth									
1823. Birth aided by drinking vegetable concoction	2	-	-	+	-	-	+	+	+
Milk Teeth									
1911. Thrown away		+	+	+	+	-	-	+	+
1916. Over house	2	+	-	+	+	-	-	-	+
Name									
1920. Given soon after birth (up to six months)	3	+	-	+	-	+	-	-	-

TABLE 4 -- Concluded

		Gal 1	Gal 2	Tol Di	Tol Du	Yur 1	Yur 2	Hup 1	Hup 2
Girl's Puberty									
1948. Covered or veiled when going outside	5	+	+	+	-	+	+	+	+
1952. Looking at people taboo	2	+	+	+	-	-	+	+	+
1961. Work compulsory, getting wood	3	-	-	-	+	+	-	-	+
1999. Men and women dance separately	2	-	-	+	-	-	-	+	-
Death									
2083. Canoe of dead broken	2	+	-	+	+	-	+	+	+
2084. House of dead burned	2	-	+	-	+	-	-	-	-
War									
2281-2. War paint		+	+	+	+	+	+	+	+
2282. Red	3	+	-	+	-	-	+	+	+
2283. Black		-	+	+	+	+	+	+	+
2284. White	2	-	-	+	-	-	+	+	+
2293. Prisoners enslaved	2	-	-	+	-	-	-	+	-
2326. Dance of incitement: abreast	2	+	-	-	+	-	-	+	+
Shamanism									
2424. Power from human spirit or ghost	2	+	-	+	-	+	+	+	+
2431. Power from reptiles	2	+	+	+	-	-	-	+	-

REFERENCES CITED

Bibliographical abbreviations used

AA	American Anthropologist.
BAE-B	Bureau of American Ethnology, Bulletins.
JRAI	Journal of the Royal Anthropological Institute (Great Britain).
UC-AR	University of California, Anthropological Records.
UC-PAAE	University of California, Publications in American Archaeology and Ethnology.

Barnett, H. G.
1937. Culture Element Distributions: VII--Oregon Coast. UC-AR 1:155-204.

Chesire, L., Saffir, M., Thurstone, L. L.
1933. Computing Diagrams for the Tetrachoric Correlation Coefficient. University of Chicago Bookstore.

Curtis, E. S.
1924. The North American Indian. Vol. 13.

Driver, H. E.
In press. Culture Element Distributions: X--Northwest California Area. UC-AR.

Drucker, Philip
1937. The Tolowa and their Southwest Oregon Kin. UC-PAAE 36:221-300.

DuBois, Cora, and Demetracopoulou, Dorothy
1932. A Study of Wintu Mythology. Journal of American Folk-lore, 45:375-500.

Goddard, Pliny Earle
1903. Life and Culture of the Hupa. UC-PAAE 1:1-88.

Kroeber, A. L.
1925. Handbook of the Indians of California. BAE-B 78.

Pearson, Karl
1900. On the Correlation of Characters Not Quantitatively Measurable. Philosophical Transactions 195A:1-47. The original derivation of r_t.
1913. On the Probable Error of a Coefficient of Correlation as Found from a Four-fold Table. Biometrika 9:22-27.

Spearman, C.
1904. The Proof and Measurement of the Association between Two Things. Amer. Jour. Psychol. 15:72-101.
1907. Demonstration of Formulae for True Measure of Correlation. Amer. Jour. Psychol. 18:161-69.

Waterman, T. T., and Kroeber, A. L.
1934. Yurok Marriages. UC-PAAE 35:1-14.