Modeling State Origins Using Cross-Cultural Data

Peter N. Peregrine, Carol R. Ember and Melvin Ember
Cross-Cultural Research 2007 41: 75
DOI: 10.1177/1069397106295445

Cross-Cultural Research
Volume 41 Number 1
February 2007 75-86

10.1177/1069397106295445
http://ccr.sagepub.com
hosted at
http://online.sagepub.com

Modeling State Origins Using Cross-Cultural Data

Peter N. Peregrine
Lawrence University
Carol R. Ember
Melvin Ember
Human Relations Area Files at Yale University

The emergence of states has been a focal concern of archaeology for many years. Although comparative methods have played a significant role in these approaches, controlled comparisons employing large samples of cases and multivariate statistical methods of analysis have been rare. The authors follow up on a previous article by employing multivariate data coded for a large sample of archaeological cases and a variety of statistical methods to empirically examine state origins.

Keywords: *cross-cultural research, cultural evolution, state origins*

The emergence of centralized polities with a hierarchy of three or more levels of decision-making above the community—what we generally call "states"—has been a focal concern of archaeology for many years. We of course realize that our definition, and indeed the term "state" itself, encompasses a wide range of variation, as the first author discussed in Blanton, Feinman, Kowalewski, and Peregrine (1996). Approaches to the question of how and why states formed have taken a wide range, and although multiple variables are often cited as being important to state origins, most approaches focus on a single variable or group of related "prime mover" variables (Service, 1975; Trigger, 1998; Yoffee, 2005). The most commonly employed "prime mover" variables are social conflict and population pressure. Social conflict explanations for state origins posit that emerging conflicts between groups with differing interests, especially involving social status or political authority, lead to internal conflict and warfare, and societies are pressured to develop hierarchical mediating structures (e.g., Engels, 1884/1972; *cf.* Claessen,

Authors' Note: The authors thank two external reviewers for their constructive suggestions.

1978). Population pressure explanations posit that humans innovate both socially and technologically in the face of resource scarcity, and such innovations (or in some cases, war or conquest) are what drive political centralization (e.g., Boserup, 1965; *cf.* Carneiro, 1970).

Multivariate approaches to state origins are rare, and controlled comparisons employing large samples of cases and statistical methods of analysis are virtually non-existent (Peregrine, 2001). In a recent article we advanced a multivariate approach to examining cultural evolution, including state origins, through the statistical technique of Guttman scaling (Peregrine, Ember, & Ember, 2004). We employed a random sample of 20 cases from the Human Relations Area Files (HRAF) Collection of Archaeology to demonstrate that a 15-item index of cultural complexity (adapted from Murdock & Provost, 1973) formed a Guttman scale. We further tested the scale order using eight evolutionary sequences taken from the *Encyclopedia of Prehistory* (Peregrine & Ember, 2001-2002). Finally, we performed a cluster analysis on 289 archaeological cases from the *Encyclopedia of Prehistory* to demonstrate that some items in the scale appear to evolve together (or at least the evolutionary distance between them is small enough so that it may not be detectible with the statistic we used—see Carneiro, 1968, p.363).

We suggested that Guttman scaling is a potent method for examining cultural evolution (Carneiro, 1963, 1968), as Guttman scales imply an evolutionary order in the way traits appear (see Table 1, A). Lower order traits on a Guttman scale are by definition precedent to higher order traits. Although it is a basic aspect of Guttman scaling, we did not specifically explore the evolutionary structure of the Guttman scale in our previous article. Here we apply the statistical technique of causal modeling to our Guttman scale variables to determine whether evolutionary relationships between these variables can be identified; that is, to determine if any one variable in the scale might be a functional prerequisite to any other.

A Causal Model of State Origins

Causal modeling employs quantitative measures to illustrate connections between variables. It allows for various assumptions about the possible directions of relationships among multiple variables to be evaluated (Birnbaum, 1981). A simple causal model based on our Guttman scale of cultural evolution (Table 1, A) would posit direct connections between each item in the scale and the next higher item; for example, the presence of ceramics (Item 1 in the scale) would be seen as causing the presence of domesticates (Item 2

Table 1
Guttman Scale of Cultural Evolution

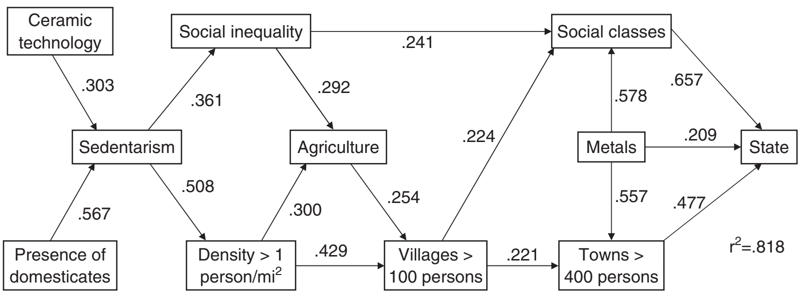
(A)	(B)
1. Presence of ceramic production	Cluster A (scale items 1 to 7) = sedentary food producers
2. Presence of domesticates	
3. Presence of sedentary communities	
4. Presence of status or wealth differences	
5. Population density > 1 person/mi^2	
6. Reliance on food production	
7. Presence of villages with > 100 persons	
8. Presence of metal production	Cluster B (scale items 8 to 11) = centralized polities
9. Presence of social classes	
10. Presence of towns with > 400 persons	
11. Presence of a political state (3+ levels of hierarchy)	
12. Population density > 25 persons/mi^2	Cluster C (scale items 12 to 15) = bureaucratic polities
13. Presence of wheeled transport	
14. Presence of writing of any kind	
15. Presence of money of any kind	

Note: Scale developed by the authors based on existing coded data for 289 archaeological cases. Variables are dichotomous. Lower numbered scale items evolve before higher numbered items. Column B identifies clusters of scale items that the authors identified as co-evolving in what might be considered "punctuated" events (Peregrine, Ember, and Ember, 2004).

in the scale), which, in turn, would be seen as causing sedentarism (Item 3). However, we demonstrated that some items in this scale evolve more-or-less simultaneously (clusters of co-evolving items are presented in Table 1, B), and this complicates any model developed from the scale, as lower scale items may be precursors to more than one higher scale item.

The Guttman scale we developed also suggests that the evolution of these traits is recursive, that is, unidirectional, allowing us to apply a relatively simple technique to identify causal connections (Asher, 1983). In this technique, called Simon-Blalock by statisticians (Asher, 1983), a model is created and partial correlation coefficients are calculated controlling for all variables identified by the model as affecting the variables being correlated. Coefficients that approach 0 or that fail to meet a minimum strength are dropped from the model. We calculated partial correlations for each pair of adjacent variables in the Guttman scale controlling for all other variables lower on the Guttman scale, and dropped marginally significant correlations (with higher than a .01 level of significance) from the model. Because

Figure 1
Causal Model for State Origins



Note: Based on the authors' established Guttman scale of cultural evolution (Peregrine, Ember, and Ember, 2004). Partial correlation coefficients are shown. Only coefficients greater than .2 are included in the model. All are significant at the .01 level.

a number of the remaining connections were weak but remained statistically significant, we simplified the model by removing all partial correlations lower than .200. Figure 1 presents the resulting model.[1]

It should be emphasized that employing the Simon-Blalock technique also allowed us to create a model without an a priori theory about the connections between the Guttman scale variables. An alternative way of using causal modeling would have been to test models with different predicted connections between variables, and then determine which provided the best fit with the data. We chose not to follow this approach, but rather to work inductively from the data themselves to see what relationships were manifest within them. Our reason is that although we believe cultural evolution is multivariate (i.e., that several causal factors are involved), existing theory tends to prioritize one "prime mover" variable over others. We hoped to avoid such prioritization to determine whether our multivariate (albeit unilineal) view was supported by the data. We think that it has been.

The model we developed implies that both domesticates and ceramic technology are precursors (or functional prerequisites) of sedentary life, and that sedentarism alone is a precursor (or a functional prerequisite) of both social inequality and increased population density. We interpret this as suggesting that sedentary life is rooted in technological changes that allow the manipulation of local foods, both in terms of the types and amounts of particular foods available and the ability to store and cook them. It is important to note that the model's implied connection between the presence of domesticates and sedentarism contradicts the established archaeological record in several parts of the world, most notably the Levant, where sedentarism begins before either domesticates or ceramics are present (Henry, 1989). We argue that these contradictory locations are unique, primarily because they are places where domestication was independently invented, and should be understood as unusual cases rather than the typical ones the model is intended to reflect.[2]

The model identifies the direct precursors to agriculture—the reliance on domesticated plants and animals—as both social inequality and population density. Although population density is not always an indication of population pressure (Ember & Ember, 1992; Roscoe, 1993), the model implies that both social inequality (defined here as the presence of status or wealth differences—see Table 1, A) and population pressure (defined here as population density greater than one person per square mile—again, see Table 1, A) are prerequisites to the transformation of people from food collectors to food producers. The model identifies a link between reliance on food production and the emergence of villages, and a link between villages and the emergence of both towns and social classes. This appears to support social conflict theories, which predict that social differences increase with village life and lead to the emergence of social classes (Diakonov, 1991).

The model identifies social classes and towns as precursors of states, but interestingly, the presence of metal technology is strongly linked to all three. Metal technology is not linked to agriculture or to villages, and thus seems to emerge without precedent in the model. That metals are a precursor of states is empirically untrue in several world areas, most notably Mesoamerica, where metals were not used until after states emerged (Blanton, Feinman, Kowalewski, & Nicholas, 1999). We posit this is because the metal technology variable represents one point on a continuum of technological sophistication that is not well captured in our Guttman scale (the other end of the continuum may be ceramic technology), and rather should be seen as representing the idea that there is a general level of technological sophistication that is a prerequisite to states. The metal technology variable is clearly problematic, and requires further examination.[3] It may be that metal technology also indicates a reliance on trade, a possibility we discuss later that may explain why metal technology is a key component of the model.

The overall fit of the model (that is, the presence or absence of a state as explained by the modeled relationships) is good ($R^2 = .818$—essentially the model accurately predicts the presence or absence of a state about 82% of the time), and we are satisfied that it captures a general, if coarse, picture of state origins. The model, however, is simply that—a model. It is not a formal test of causal relationships and cannot prove that change in one variable

Table 2
Means for Precedent Cases and Sample Cases

Variable	Precedent Mean	Sample Mean	*t*-value (*df* = 28)
Ceramic technology	.9	1.0	–1.797
Presence of domesticates	.93	1.0	–1.440
Sedentarism	.86	.97	–1.797
Social inequality	.9	1.0	–1.797
Density >1 person/mi^2	.83	1.0	–2.415
Agriculture	.83	.97	–2.117
Villages >100 persons	.69	1.0	–3.550*
Social classes	.0	.97	–28.000*
Metals	.21	.86	–7.294*
Towns >400 persons	.28	.79	–5.477*
State	0	1.0	

Note: Means are for all sample cases where a state evolved, with *t*-values of mean differences.
*Significant at the .001 level

causes change in another. With at least two problematic variables (ceramics and metals), it seems reasonable to ask whether the model identifies true relationships. Because our data are diachronic, they provide the opportunity to independently examine one definitive aspect of causal relationships—that change in a presumed causal variable precedes presumed effects (Ember & Ember, 2001).

Testing the Causal Model

To test whether change in presumed causal variables preceded presumed effects, the first author identified ancestral cases for each case in the sample, and used the ancestral case to code each sample case for its precedent value.[4] Table 2 presents mean differences between precedent case values and sample case values for all sample cases where a state evolved (that is, the ancestral case did not have a state but the case itself did). It is clear from Table 2 that presumed causal variables show significant change between the precedent case and sample case values, suggesting that the causal model does identify causal relationships at least coarsely. Specifically, the three variables identified in the model as precursors of states show very large mean differences (and correspondingly large *t*-values). Variables not identified as direct precursors of states show small and non-significant mean differences. Table 2 provides strong evidence that the variables identified as causal in the model do undergo statistically significant change when states evolve.

Another way to test the value of the model is to determine whether the variables identified as direct precursors of states accurately predict variation in political integration. This requires a somewhat different set of data from that used in the causal model and *t*-tests. First, a set of variables capturing variation beyond simple binary differences is desirable. Second, because the processes of variation predicted here are temporal—states develop over time out of non-states—time series data are also desirable. The data set used both in the causal model and the earlier Guttman scale includes a number of ordinal time series variables, and we were able to use those variables here (Peregrine, 2003). We chose four variables for time series regression: political integration as the dependent variable, and urbanization, social stratification, and technological specialization as the independent variables.[5]

Time series regression differs from other forms of regression as a temporal element is included in the regression equation. Because there is often time-dependent autocorrelation between variables (the variables tend to change over time independent of the causal relations among them), a transformation of the variables is usually done to minimize autocorrelation effects. Here we employ the Cochrane-Orcutt transformation, which is simply an algorithm used to correct data for autocorrelation effects, and provide the value of the Durban-Watson *d* statistic to estimate the extent of autocorrelation following transformation (Ostrom, 1990). The Durban Watson statistic is a test of whether the model errors are independent. If they are not, then autocorrelation is present, and the resulting regression coefficient may be artificially inflated.

The regression values presented in Table 3 demonstrate that the variables identified as direct precursors of states in the model are good predictors of variation in political integration. Each row represents a different regression model and each predicts variation in political integration over time with better than 90% accuracy. The first three rows show the regression results with each variable as a single predictor. The next three rows show two variables as predictors and the last row all three variables as predictors. In combination, the variables (two at a time or all three) do a better job of predicting variation, but the improvement is only on the order of 3-5% over the individual variables. This implies that these variables all share a significant link to the origin of states and, therefore, that one or more underlying factors tie these three variables together. The question is, what might this factor or factors be? We discuss two possibilities—economic vulnerability and scalar stress.

Table 3
Time Series Regression Coefficients and Durban-Watson *d* Statistics for Various Models Predicting Variation in Political Integration.

Variables Used to Predict Variation in Political Integration	R^2	*d*
Social stratification (S)	.988	0.7[a]
Technological specialization (T)	.944	0.8[a]
Urbanization (U)	.938	2.0
U, S	.966	0.7[a]
U, T	.990	1.8
S, T	.985	0.7[a]
U, S, T	.982	0.7[a]

a. May indicate that autocorrelation effects remain.

Economic Vulnerability

The archaeological record suggests that the earliest metalworking was for personal ornaments and weapons (see Note 3). Because ores are rarely found in every community, extensive metalworking probably implies both the presence of trade at considerable distance from home and the presence of markets to exchange goods. Goods and individuals moving along trade routes would need to be protected, so military defenses may also increase. Protection of longer-distance routes would require some multi-community solution, either through alliances or more formalized political hierarchies (as suggested by M. Ember, 1968; see also Wright & Johnson, 1975, for discussion of the role of longer distance trade). Although metallurgy is the only variable in our model clearly suggestive of longer distance trade, we suggest that such trade is increasingly likely with increasing sedentarization as there may be less self-sufficiency locally and more demand for produce, timber, and other goods (e.g., metals) from greater distances. Communities that are not self-sufficient in food production would be particularly vulnerable to shortages caused by wars that curtailed their supplies from the hinterland. With this increasing dependence on longer distance trade comes increasing vulnerability to disruption by other political groups. It is this vulnerability that may foster a more centralized political military response (M. Ember, 1968).

Scalar Stress

Based on small-group and organizational research, Johnson (1982) argues that decision making suffers and conflict increases when decisions have to be made by more than six functional entities (individuals, families, departments, etc.). Johnson calls this *scalar stress*, and he describes two major responses to solve the problems associated with such stress. The first, common to egalitarian societies, is a horizontal strategy of fissioning. When a community gets too large for effective decision-making, it may split into two or more independent units. Another response is vertical—a hierarchy is established with superordinate decision-making. So, if families are organized into higher-level units (e.g., neighborhood or kin groups), these higher-level units could effectively serve to make decisions as long as there were no more than six such units. As for why a vertical solution would be chosen over a horizontal solution, Johnson suggests that some conditions, such as environmental circumscription (following Carneiro, 1970), make fission not a viable strategy. Sedentarism with permanent dwellings, agricultural features (e.g., irrigation, terraces), and investment in specialized equipment may also make fission a less likely strategy (Bandy, 2004). However, as we have just described it, scalar stress does not clearly suggest why towns with more than 400 persons, metals, and social classes would lead to the hierarchy of three or more decision-making units above the community that we call the state. Towns of more than 400 would clearly need one level of hierarchy beyond the family to function well, but why higher? We suggest that it may be the emergence of an elite that gives rise to integrative multicommunity mechanisms. After all, there are relatively few elites in any community, and therefore there would be interest among the elite in reaching out to other elites to marry, trade with, and solidify rules and regulations that would protect accumulated wealth, etc. (as examined by Helms, 1979). The emergence of personal ornaments that are difficult to acquire suggests that the elite would have to expend a lot of resources to justify and solidify their position. As soon as vertical leadership developed, we would argue that there is little to stop the accumulation and control of resources by leaders, particularly when leaders begin to reach beyond the borders of their communities for support from other elites. In addition, if we add the effect of economic vulnerability that we discussed above, there would be a considerable number of functions that leaders might be handling—longer-distance trade, competition among groups over trade routes, markets and their regulation, and judicial functions to handle attempts to gain access to wealth by theft or rebellion. Johnson also points out that leaders cannot handle that many functions efficiently, and in conjunction with the number of units may press toward the development of more hierarchy (see Lee, 2004 for an archaeological case study).

Conclusion

We want to conclude by returning to the original subject of this article—multivariate analysis. In this and our previous article (Peregrine, Ember, & Ember, 2004), we have illustrated multivariate approaches to understanding complex phenomena in human prehistory. We think these articles demonstrate the value of such approaches. Whereas most existing work on complex phenomena like state origins focus on a single prime mover variable, our causal model and tests of causality imply that no simple prime movers exist. Rather, our work suggests that at minimum we must seek complex factors such as economic vulnerability combined with conditions leading to vertical solutions to scalar stress to understand a phenomenon like state origins. Cultural evolution appears to be multivariate and multicausal, and multivariate statistical techniques should be employed in research on cultural evolution, because simple univariate models and analyses do not capture the multivariate processes at work.

Notes

1. It should be noted that only eleven of the fifteen scale items are included in the causal model. This is because the Guttman scale identified four of the variables as evolving after states, and they were therefore not included.

2. The causal model is based on statistical tendencies, and there will always be cases that do not fit. Because the model is one of an evolutionary process, it may be logical that the initial cases to evolve follow more diverse paths than later cases (i.e., that evolution "tests" several paths before "selecting" a typical one).

3. It is interesting to note that metals are often initially used as personal ornaments and weapons. Personal ornaments as markers of status or office suggest an obvious connection between metals and states, but not necessarily a causal one. Weapons, on the other hand, suggest connections between warfare or conquest and states, a connection that has a strong theoretical and empirical foundation (e.g. Carneiro, 1970).

4. The data set employed in these analyses is the same used by Peregrine, Ember, and Ember (2004) to develop the Guttman scale and has been published in full by Peregrine (2003). They consist of 289 archaeological cases ranging in time from roughly 2 million years ago to roughly 500 years ago. The cases are coded for 35 variables, including a 10-variable cultural complexity scale developed by Murdock and Provost (1973). The first author recoded these 10 variables into the 15 dichotomous variables used here, as described in Peregrine, Ember, and Ember (2004). Ancestral cases were identified using the "Table of Descent" in Peregrine and Ember (2002).

5. Values for variables used in the time series regression are as follows:

Political integration

1 = Autonomous local communities
2 = 1 or 2 levels above community
3 = 3 or more levels above community

Social stratification
- 1 = Egalitarian
- 2 = 2 social classes
- 3 = 3 or more social classes or castes

Technological specialization
- 1 = None
- 2 = Pottery
- 3 = Metalwork (alloys, forging, casting)

Urbanization (largest settlement)
- 1 = Fewer than 100 persons
- 2 = 100-399 persons
- 3 = 400+ persons

References

Asher, H. (1983). *Causal Modeling*. Thousand Oaks, CA: Sage.

Bandy, M. S. (2004). Fissioning, Scalar Stress, and Social Evolution in Early Village Societies. *American Anthropologist*, *106*(2), 322-333.

Birnbaum, I. (1981). *Introduction to Causal Analysis in Sociology*. London: MacMillan.

Blanton, R. E., Feinman, G., Kowalewski, S., and Nicholas, L. (1999). *Ancient Oaxaca*. New York: Cambridge University Press.

Blanton, R. E., Feinman, G., Kowalewski, S., Peregrine, P. (1996). A Dual-Processual Theory for the Evolution of Mesoamerican Civilization. *Current Anthropology*, *37*, 1-14.

Boserup, E. (1965). *The Conditions of Agricultural Growth*. Chicago: Aldine.

Carneiro, R. (1963). Scale Analysis as an Instrument for the Study of Cultural Evolution. *Southwestern Journal of Anthropology*, *18*, 149–169.

Carneiro, R. (1968). Ascertaining, Testing, and Interpreting Sequences of Cultural Development. *Southwestern Journal of Anthropology*, *24*, 354-374.

Carneiro, R. (1970). A Theory of the Origin of the State. *Science*, *169*, 733-738.

Claessen, H. J. M. (1978). The Early State: A Structural Approach. In H. J. M. Claessen and P. Skalnik (Eds.), *The Early State* (pp. 533-596). The Hague, The Netherlands: Mouton.

Diakonov, I. M. (1991). *Early Antiquity*. Chicago: University of Chicago Press.

Ember, C. R. and Ember, M. (1992). Resource Unpredictability, Mistrust, and War: A Cross-Cultural Study. *Journal of Conflict Resolution 36*, 242-262.

Ember, C. R. and Ember, M. (2001). *Cross-Cultural Methods*. Walnut Creek, CA: AltaMira.

Ember, M. (1968, November). *The Social Environment and the Rise of Civilization*. Paper presented at the American Anthropological Association, Seattle, WA.

Engels, F. (1972). *Origin of the Family, Private Property, and the State*. New York: International Publishers. (Original work published in 1884)

Helms, M. (1979). *Ancient Panama: Chiefs in Search of Power*. Austin: University of Texas Press.

Henry, D. O. (1989). *From Foraging to Agriculture: The Levant at the End of the Ice Age*. Philadelphia: University of Pennsylvania Press.

Johnson, G. A. (1982). Organizational Structure and Scalar Stress. In C. Renfew, M. J. Rowlands, and B. A. Segraves *(Eds.), Theory and Explanation in Archaeology*, (pp. 389-421). New York: Academic.

Lee, Y. K. (2004). Control Strategies and Polity Competition in the Lower Yi-Luo Valley, North China. *Journal of Anthropological Archaeology 23*(2), 172-195.

Murdock, G. P. and Provost, C. (1973). The Measurement of Cultural Complexity. *Ethnology 12*, 379-392.

Ostrom, C. (1990). *Time Series Analysis: Regression Techniques*. Thousand Oaks, CA: Sage.

Peregrine, P. N. (2001). Cross-cultural Comparative Approaches in Archaeology. *Annual Review of Anthropology 30*, 1-18.

Peregrine, P. N. (2003). Atlas of Cultural Evolution. *World Cultures 14*(1),1-89.

Peregrine, P. N., Ember, C. R., and Ember, M. (2004). Universal Patterns in Cultural Evolution: An Empirical Analysis Using Guttman Scaling. *American Anthropologist 106* (1), 145-149.

Peregrine, P. N. and Ember, M. (Eds.). (2001-2002). *Encyclopedia of Prehistory* (9 Vols.). New York: Kluwer/Plenum.

Roscoe, P. (1993). Practice and Political Centralization: A New Approach to Political Evolution. *Current Anthropology 34*(2), 111-140.

Service, E. (1975). *Origins of the State and Civilization: The Process of Cultural Evolution*. New York: W.W. Norton.

Trigger, B. G. (1998). *Sociocultural Evolution*. Oxford: Blackwell.

Wright, H. and Johnson, G. (1975). Population, Exchange and Early State Formation in Southwestern Iran. *American Anthropologist 77*, 267-289.

Yoffee, N. (2005). *Myths of the Archaic State*. Cambridge: Cambridge University Press.

Peter N. Peregrine is Professor and Chair of the Department of Anthropology at Lawrence University in Appleton, Wisconsin. His research focuses on cultural evolution and cross-cultural methods. He has done archaeological research on the origins of complex societies in North America, Europe, and Asia. His most recent work has focused on cross-cultural approaches to the study of unilineal evolution.

Carol R. Ember is Executive Director of the Human Relations Area Files at Yale University. She has served as President of the Society for Cross-Cultural Research and co-directed the Summer Institutes for Comparative Anthropological Research, which were supported by the National Science Foundation. Her textbook, *Cross-Cultural Research Methods*, written with Melvin Ember, won a Choice award as an outstanding academic title. Most of her research career has been devoted to cross-cultural research on variation in marriage, family, kin groups, gender roles, predictors of war, and other forms of violence.

Melvin Ember is President of the Human Relations Area Files at Yale University. His cross-cultural studies—on marriage, family, kin groups, political evolution, linguistics, war and peace, and other aspects of cultural variation—have been supported by the National Science Foundation and the U.S. Institute of Peace, among others. He has edited or co-edited various multi-volume encyclopedias, including *Countries and Their Cultures* and the *Encyclopedia of Urban Cultures,* and is the co-author of various textbooks in anthropology, including *Cultural Anthropology*, now in its 12th edition.