Cultural Mixture Modeling: Identifying Cultural Consensus (and Disagreement) using Finite Mixture Modeling

Shane T. Mueller (smueller@ara.com) and **Elizabeth S. Veinott (bveinott@ara.com)**
Cognitive Science Group, Klein Associates Division, ARA Inc., Fairborn OH 45324 USA

Abstract

In this paper, we describe a new technique for identifying cultural consensus called Cultural Mixture Modeling (CMM). This technique adopts finite mixture modeling, and introduces a new probabilistic formulation of agreement, which we call the *strong consensus model*. We use this technique to examine the cultural belief data from Weller (1983; 1984) and social network data from Krackhardt (1987). We show that CMM can go beyond classic models of consensus and identify situations in which multiple distinct but disagreeing beliefs exist between subgroups of individuals. By identifying groups of shared belief, CMM offers a practical and useful technique for understanding and characterizing how socio-cultural factors influence our beliefs and attitudes.

Keywords: culture, mental models, consensus theory

Background

Understanding the underlying beliefs, attitudes, and mental models of individuals is an important goal in a number of domains of cognitive science. This is important for applied problems (in which these mental models might be elicited in order to develop training, design system interfaces, or understand a target population), as well as basic research problems (e.g., identifying a concept's conceptual coherence; determining typical associations from verbal stimuli). We have found it is especially useful when studying how those beliefs and mental models are affected by social or cultural factors, and identifying how different beliefs lead to different behaviors. For example, one sociological view of culture (cf. Atran, Medin & Ross, 2005; Sieck, Smith & McHugh, 2007) holds that culture is comprised primarily of the shared beliefs and practices of a group, rather than just the demographic and linguistic characteristics commonly equated with culture. A pernicious problem faced when eliciting such knowledge is in knowing whether variation among respondents simply represents random noise, or whether that variation represents some more fundamental differences in what a group of individuals believe.

One method that has been developed to understand whether a group of people share a set of common beliefs is called Cultural Consensus Theory (CCT; Romney, Batchelder, & Weller, 1986). CCT is a set of statistical tools designed to assess agreement in belief or knowledge among a set of respondents. Perhaps CCT's most profound insight is that culturally-correct responses can be determined "without the answer key": the culturally-correct beliefs are the ones that most members of that culture consistently agree with. CCT uses a matrix-algebra procedure known as eigenfactor decomposition to determine whether or not a consensus exists. This procedure starts by forming a dissimilarity matrix across respondents, and then decomposes the matrix into its principle components, thus determining whether a consensus exists among the respondents. In essence, CCT is similar to factor analysis performed on the responses of a survey, but instead of determining sets of questions for which respondents give similar responses (i.e., the columns), it determines sets of respondents who share similar beliefs (i.e., the rows). If the respondents are well described by a single factor, then a consensus is deemed to exist.

If a consensus does exist, one can estimate the extent to which each respondent agrees with the dominant belief set. Romney, Batchelder and Weller (1986) refer to this agreement as cultural competence. Cultural competence has often been found to be related to demographic factors such as age (for example, older and more experienced individuals are more likely to believe the culturally correct answer). Thus, by determining the culturally-correct answers, CCT allows each individual to be given a score showing how well they know those answers.

Limitations of Cultural Consensus Theory

Although CCT has proven useful in understanding whether respondents in a survey or interview share common beliefs, it is not without its limitations. The most obvious limitation is that the model only determines whether or not an overall consensus exists, but not whether there are multiple subcultures who believe different things. If a consensus does not exist, there are several plausible explanations that CCT cannot distinguish between. One possibility is that there is no consensus because each respondent is essentially unique. Another possibility is that there are several subsets of consistent beliefs. As an illustration (expanded in Demonstration 1), consider using this method to understand the positions of U.S. politicians. Across a political body (such as the U.S. Senate), a consensus would be unlikely. However, lack of consensus does not mean that each Senator's response patterns are completely unique: we would likely find a handful of coherent beliefs aligned with political party membership and geographic region. CCT can determine whether members agree, but if they do not agree, it is incapable of providing much insight without placing a priori beliefs about what the groups should be (e.g., political affiliation). But in that case, CCT may not be necessary; we can simply compare the range of responses for each pre-defined sub-group and determine whether they differ.

The insights of CCT have proven useful, but some of its restrictions are difficult to surmount in principled ways. To address some of these problems, we have adapted a statistical technique called finite mixture modeling (FMM; Leisch,

Table 1: Comparison of Cultural Consensus Theory and Cultural Mixture Modeling.

Dimension	Cultural Consensus Theory	Cultural Mixture Modeling
Common Truth Assumption	There is a single fixed "answer key" all respondents adhere to.	Multiple "answer keys" can exist.
Error Variance	Errors are conditionally independent.	Covariance modeled explicitly.
Response Items	Responses are recoded into dissimilarity or correlation matrices using heuristics.	Responses are explained explicitly with a generative model tailored to response type.
Respondent Competency	Each respondent has a fixed competence.	Each respondent has a measurable competence (likelihood) for each identified group.
Statistical Procedure	MLE Factor Analysis using eigen decomposition.	Finite mixture modeling using E-M optimization and BIC.
Results	Whether a consensus exists; competencies of respondents.	Number of beliefs, response patterns of each group, competencies of respondents.
Statistical Inference	Rules of thumb: large 1st eigenvalue; 1st/2nd eigenvalue $>$ 3.0; loadings $>$ 0.	Probabilistic reasoning using BIC and maximum likelihood.

2004) to identify groups of shared belief. Instead of asking whether or not a consensus exists, our method (called Cultural Mixture Modeling: CMM) determines the optimal number of groups of shared belief that generated the observed data. Thus, in order to determine whether a consensus exists, it identifies how many differing sets of beliefs exist, and what those beliefs are. Both CCT and CMM are similar in that they look for consistent patterns of belief across a complete set of questions. But CMM has a number of differences from CCT, which are described briefly in Table 1.

Some of the differences between CCT and CMM arise because the developers of CCT used computational techniques that were available and well-understood at the time. In contrast, CMM uses techniques that have become more widely used in the intervening years, especially in the field of advertising for market segmentation (Leisch, 2004). CMM uses the E-M algorithm (Dempster, Laird, & Rubin, 1977) on the complete *NxM* response set (with *N* respondents and *M* responses), which typically requires more computation and computer memory storage than eigen decomposition, which requires inverting an *NxN* matrix.

Cultural Mixture Modeling and the Strong Consensus Model for binary agreement data

The statistical theory involved in finite mixture modeling is fairly well-developed and understood, and multiple free and commercial software packages exist that allow fairly complex models to be developed and applied. It is important to note that CMM does not use bayesian inference algorithms to compute its solutions, although it does frame the problem in the probabilistic language of a generative model, and uses the Bayesian information criterion (BIC) to select the simplest most descriptive model.

CMM begins by identifying a probabilistic model that generated the responses. The data sets analyzed in this report involved binary responses, in which one value is coded as 0 and a second value 1. Such responses can be described by a binomial probability model, in which a parameter γ determines the probability that a member of group i makes response 1 on question j.

$$p_{i,j}(x_{i,j}) = \gamma^{x}_{i,j}(1-\gamma)^{(1-x_{i,j})} \quad (1)$$

The interpretation of γ is important from both psychological and statistical perspectives. One possible interpretation views γ as the tendency or certainty of an individual to make the response coded as "1". In this case, someone may be equivocal ($\gamma = .5$); certain ($\gamma = 0$ or $\gamma = 1$); or leaning toward one response (e.g., $\gamma = .8$). Here, one assumes that the value of γ is a real micro-cognitive value which leads to one response on one occasion, and the other response on another occasion, for a given individual. Such an interpretation may be appropriate for questions on which respondents do not hold well-conceived opinions (e.g., "Should bicycle helmets be certified by state or federal agencies?"), for questions the respondents are likely to be guessing ("Is Moscow further north than Minneapolis?"), for questions in which true beliefs are transitory ("Are you tired?") or are likely to change from situation to situation (e.g., "Do you want to eat pancakes for breakfast?"). In these cases, when a group of respondents is identified, γ for each response is set to be equal to the proportion of respondents giving the affirmative response.

In contrast, some beliefs are not characterized by strength-like tendencies. In many cases, the response a person gives is likely to be well-conceived and unlikely to easily change because of the context or situation (e.g., "Are you a Democrat or Republican?"). In these cases, the interpretation of γ as a tendency is inappropriate, because the proportional make-up of a group may not reflect the belief strength of individuals. For example, on some set of questions about religious belief, CMM may identify a group with 18 Catholic respondents and two Lutherans. The proportion of Catholics may be .9, but this does not mean that the each respondent was 90% Catholic.

In such cases, we interpret γ as a measure of the likelihood that a group belief was adopted by the individual, a tolerance for divergence from consensus. Because the notion of tolerance should not be dependent on specific questions, we restrict γ to take on one of two values: either α or $1-\alpha$, where α is a value close to 0 (usually around .05), and is identical for all responses and respondents. We call this the *strong consensus model*.

In the strong consensus model, α describes the degree to which group membership becomes less likely for an individual, for each response that does not conform to the group's model. In its extreme (as $\alpha \rightarrow 0$), no statistical tools would be needed to fit this model: respondents would belong to the same group if they responded identically to all other member of that group. When α is non-zero, it roughly reflects number of respondents on each question that can disagree with the group consensus.

Once a probabilistic generative model has been specified, the application of CMM is fairly straightforward, using the E-M algorithm. We first specify a fixed number of groups to consider (usually starting at 1, and then increasing to a number fewer than the number of respondents). The E-M algorithm begins by randomly assigning persons to groups, computing the most likely responses according to those assignments, then re-assigning members to the group they were most likely to have come from. A value of γ (for the binomial model) or α (for the strong consensus model) is computed for each question and each group. After multiple cycles of this process, the algorithm converges to a local likelihood maxima, and by starting from multiple initial configurations fairly stable solutions can be obtained.

Models with more groups have more parameters, which tends to improve the ability to account for data. We use the Bayesian Information Criterion metric (BIC; Schwartz, 1977) to optimally counteract increases in goodness-of-fit with increases in model complexity. We used the flexmix package (Leisch, 2004) in the R statistical computing language (R core development team, 2007), which handles much of this process automatically.

In the remainder of this paper, we will describe how we have applied CMM to several problems related to cultural consensus analysis. We will begin with an illustrative example of voting records in the U.S. Senate, and then move on to more complex situations in which the solutions are less obvious. These examples will show how traditional methods of identifying consensus failed to reveal the true structure of the shared beliefs amongst respondents.

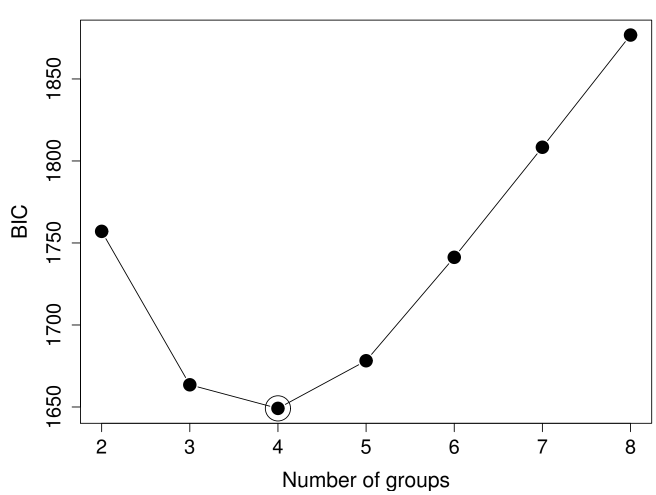
Example 1: U.S. Senate Voting Record on AFL-CIO Issues

Our first example applying considers the voting record of members of the US Senate during 2005–2006 on 19 votes that AFL-CIO leadership identified as being of interest to its members. Although senate voting patterns do not represent mental models of beliefs per se, these data help illustrate how CMM can applied and provide validation of its inferences, because there are known groups (i.e., Democrats and Republicans) with strong shared beliefs.

We first analyzed the data using the CCT approach: we computed agreement scores between Senators, performed eigen decomposition, and examined the results. Although the first factor accounted for 92% of the variance, which was 25 times the next factor, the competence loadings of more than half of the respondents were negative, suggesting that no consensus opinion existed among the Senators' votes. The basic CCT inference is that there is no consensus. At this point, a standard approach might be to divide the Senators into known 'cultural' groups (e.g., conservative versus liberal, Democrat versus Republican, Blue state versus Red state, etc.), and apply CCT to each individual group. However, CMM does not require this ad hoc exploration, as it produces the groups of shared belief (i.e., cultural groups) as an outcome of its inference process.

To apply CMM, we used the strong consensus model with $\alpha = .05$. Our reasoning is that for voting patterns, Senators will deliberate, consult advisers, and come to a conscious decision on how to vote. We computed the BIC score of the best-fitting model for each number of clusters (see Figure 1), and then selected the model with the smallest BIC score, which happened to have four groups. Models with fewer groups did not account for the data as well, and models with more groups did not increase predictability enough to counteract the increased complexity of the statistical models (which added 19 parameters for each group).

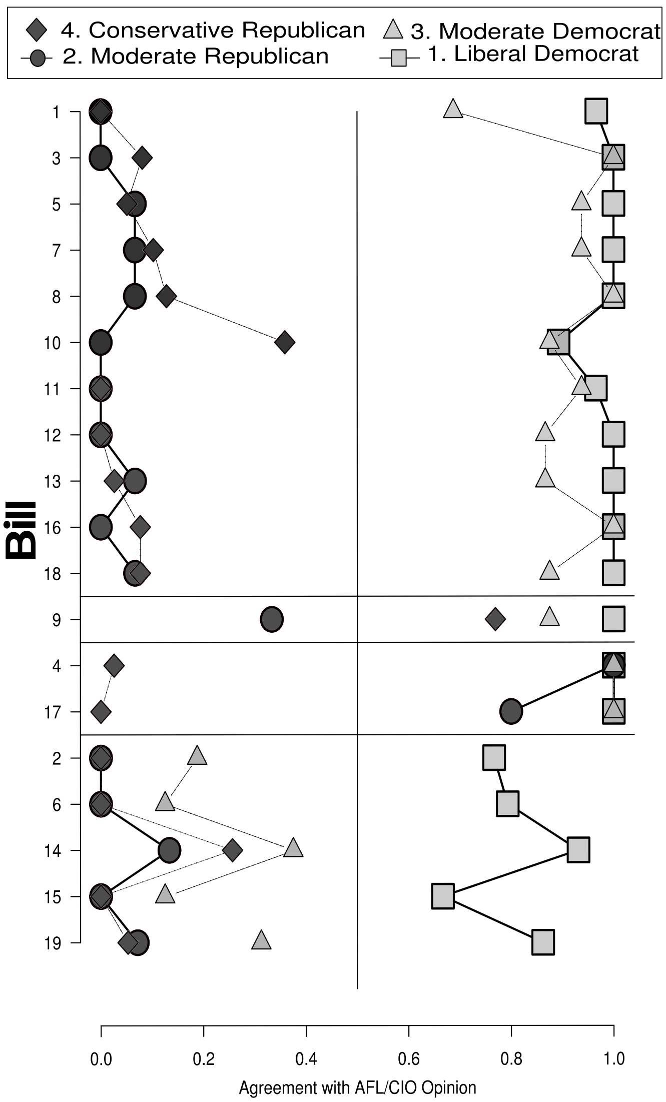
Figure 1: BIC scores for Example 1 show that the best-fitting least-complex model had four groups.



Although party membership was not used in the analysis, CMM accurately segmented along party lines, with two groups that were primarily Democrats, and two groups that were primarily Republicans. The distribution of party membership across identified groups is shown in Table 2.

Figure 2 examines how the four groups responded across

Figure 2: Proportion of Senators within each group voting in agreement with AFL-CIO. On the 11 issues at the top of the figure, the Republican groups agreed but disagreed with both Democratic groups. The two Republican groups differed on three issues (9, 4, and 17), and the two Democratic groups differed on five issues (2, 6, 14, 15, and 19).



the range of votes. Although the consensus response for each group was represented as either "Yea" or "Nay", we show the actual proportion agreeing the AFL/CIO position in Figure 2. Group 1, which consisted solely of Democrats, voted pro-union on every issue. Similarly, Group 2, which consisted solely of Republicans, voted anti-union on all issues but one. Each of the remaining two groups also identified with a single party, but demonstrated some willingness to oppose the majority of their party. Further examination (cf. Mueller et al., 2007) determined that the smaller Republican group consisted of conservative Republicans, and the smaller Democratic group consisted primarily of moderate democrats.

Table 2: Distribution of party membership across groups.

Group	Democrat	Independent	Republican
1 (Liberal Dem.)	30	0	0
2 (Mod. Rep.)	0	0	39
3 (Mod. Dem.)	14	1	1
4 (Cons. Rep)	0	0	15

This example illustrates how one of the weakness of CCT can be overcome using CMM. Standard application of CCT accurately indicated that no consensus existed among Senators on these issue, but failed to identify whether there was agreement among sub-groups. In fact, we determined there were four groups, which would have been difficult to identify from party affiliation alone. CCT may have been able to infer that consensus did not exist within parties, but in many situations, we do not know the groups the respondents should fall into, or the groups we have information about might be misleading. CMM solves this problem by allowing us to infer the cultural groups from the data. Next, we will show how this can lead to conclusions that CCT failed to make.

Example 2: Belief about Diseases

Romney (1999) used a classic anthropological data set by Weller (1983; 1984) to demonstrate the effectiveness of CCT. The data dealt with the cultural beliefs of 24 Guatemalan women about the causes and treatments of 27 diseases. Respondents were asked two questions: is the disease contagious, and is the disease treated with hot (versus cold) treatments.

We used CMM and the binomial agreement model to examine these same data. For the contagion data, CCT suggested that a consensus belief existed. As shown in Figure 3, CMM came to the same conclusion, identifying that a single cultural group best accounted for the data. In contrast, for the hot/cold treatments, CCT had concluded that no consensus existed, suggesting that there may be no consistency across the respondents. Here, CMM identified two cultural groups, (circles in Figure 3), indicating that two distinct consistent opinions existed.

To determine whether these responses corresponded to their concomitant responses on the contagion rating question, we separated the diseases into two sets: one for which the two groups agreed upon the treatment, and a second for which the two groups disagreed about the treatment. We then ordered these by their ratings on the contagion question. The results are shown in Figure 4.

Figure 4 shows that for diseases in which the two groups tended to agree upon the treatment, there was little relationship between contagion and the treatment ($r = .325$ and $r = .327$ for Groups 1 and 2, $p[r \neq 0] > .1$). However, for those in which the groups disagreed about the treatment, there were strong relationships within the two groups, in opposite directions. Group 1 tended to believe that contagious diseases should be treated with hot treatments and non-contagious

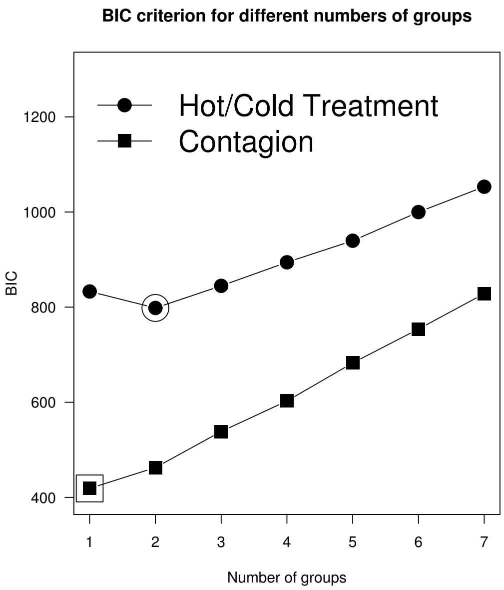


Figure 3: The BIC statistics computed for both data sets, for between 1 and 7 groups. Results show substantial agreement among all respondents for contagion data, but treatment data were best described by two distinct groups.

diseases should be treated with cold treatments ($r = .79$; $t(12) = 4.5$, $p < .001$), while Group 2 tended to believe the opposite($r = -.85$;$t(12) = 5.6$, $p < .001$). Apparently, for a certain set of diseases, there is an agreed-upon treatment across the culture. For others, there is no agreed treatment, and in these cases people adopt a belief that the treatment (hot or cold) should depend upon whether or not the disease is contagious, with the two groups adopting opposite beliefs. Post-hoc analyses showed that group membership was also related to age and number of children, so this could be an effect of age or experience.

Analyzing Weller's (1983, 1984) classic data set showed an interesting set of beliefs that standard CCT failed to reveal. Although there was no strong agreement among respondents about treatment, this did not mean that there was no agreement at all. In fact, there was agreement about treatment for about half of the diseases, and for the other half, there was strong disagreement.

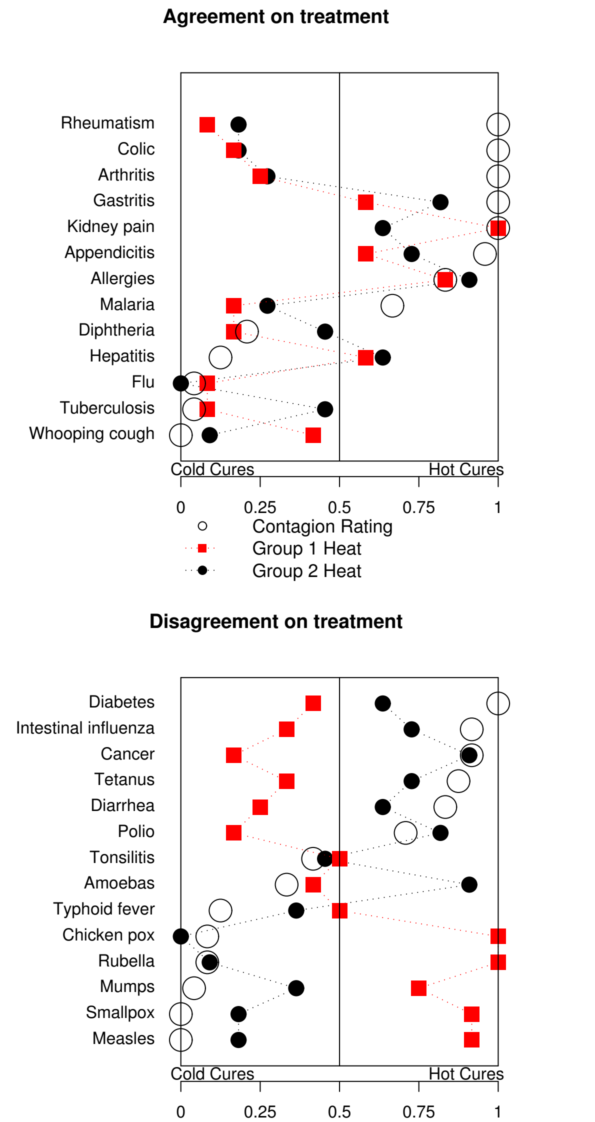


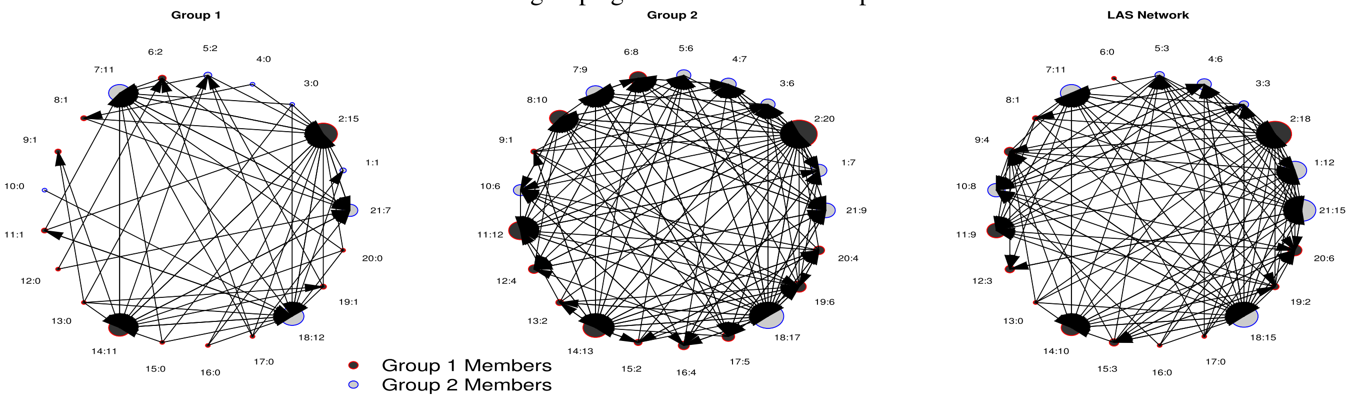
Figure 4: Mean probabilities for treatment responses, sorted by mean probability for the contagion responses. When groups agreed on the treatment (top), responses were independent of their contagion responses; when groups disagreed (bottom), responses depended on contagion.

Example 3: Cognitive Social Structures

Mental models sometimes can be inferred through detailed questionnaires (as in Example 2). Other times, the relations are more complex, and may require the use of more complex network structures to represent statements about complex relationships. One such type of mental model is called a "Cognitive Social Structure" (Krackhardt, 1987). These structures are similar to social networks (which for a set of individuals determine social relationships between individuals), but instead of asking each individual only about the relationships they are part of, it asks each individual about every pairwise relationship among the members of the group. This results in N NxN matrices, and is ideal for CMM analysis.

In Krackhardt's (1987) classic paper which introduced the concept of cognitive social structures, responses were collected data from 21 managers within a business organization. Each manager was asked about the advice relationships among the 21 managers. (A total of 441 possible connections between people.) In an attempt to identify the consen-

Figure 5: Depictions of the two distinct groups inferred from beliefs about advice relationships in Krackhardt (1987). Lines indicate that more than 50% of the members of a group agree about a relationship.



sus network, Krackhardt (1987) cited Romney, Weller and Batchelder (1986). However, they did not actually perform CCT on the data, and simply examined the average network.

We examined the data using CMM and found that, in contrast to the assumptions of Krackhardt (1987), two distinct cultural groups existed. One group (the "hierarchical" group with roughly 2/3 of the members), believed that a few strong advice relationships existed along the organization hierarchy (a few high-level managers and vice presidents tended to be sought for advice). The other group (the "democratic" group, consisting of about 1/3 of the managers) believed that advice was sought much more equally and in a more distributed fashion. The hierarchical group appears to contain the more peripheral managers, whereas the democratic group tends to contain the more central members of the team, including several of the most central managers identified by the hierarchical group. The two consensus networks are depicted in Figure 7, along with the "locally-aggregated" (LAS) network, which contained only relations which a respondent was personally part of.

This analysis suggests that a minority of more central managers tended to believe in a more democratic set of advice relationships, which was in fact fairly accurate (compared to the ground-truth LAS network). The majority of the managers (who tended to be more peripheral) believed that advice was reserved for a few central members. This suggests that the true role of the central managers in this network was not as primary sources of advice, but rather as the members of the organization who best understood how the team operated.

Discussion

In this paper, we introduced CMM and the strong consensus model, and showed how they can be used to evaluate whether a group of respondents share similar views about a set of topics. CMM works by identifying groups of respondents who share common beliefs, and provides additional flexibility over CCT because it models the individual response process, rather than just the similarity space across people. In contrast to traditional cross-cultural approaches, CMM treats culture as an outcome, revealing the nature of culture rather than taking nationality as a proxy for culture and treating it as an independent variable. This allows the nature of culture to be revealed without relying on a priori notions of how it is organized.

Acknowledgments

Research reported here was supported in part by the U. S. Army Research Laboratory through ADACTA under cooperative agreement DAAD-19-01-2-0009 (07TA2-SP1-RT1). We also thank Dan Horn and Winston Sieck for useful and helpful comments.

References

Atran, S., Medin, D. L., & Ross, N. O., (2005). The Cultural Mind: Environmental decision making and cultural modeling within and across populations. *Psychological Review, 112*, 744-776.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

Leisch. F. (2004). FlexMix: A general framework for finite mixture models and latent class regression in R. *Journal of Statistical Software, 11,* 8.

Krackhardt D. (1987) Cognitive Social Structures. *Social Networks, 9*, 109–134.

Mueller, S. T., Sieck, W. S., & Veinott E. S. (2007). Cultural Metrics: A finite mixture models approach. *Final report for the ADA/CTA Program*.

R Development Core Team (2007). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,*

Romney, A. K. (1999). Culture consensus as a statistical model. *Current Anthropology, 40 (Supplement)*, S103–S115.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist, 88*, 313–338.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*, 461–464.

Sieck, W. R., Smith, J., & McHugh, A. P. (2007, Oct. 1-5). Cross-national comparison of team competency values. *Paper presented at the 51st HFES, Baltimore, MD.*

Weller, S. C. (1983). New data on intracultural variability: The hot/cold concept of medicine and illness. *Human Organization 42,* 249–257.

Weller, S. C. (1984). Cross-cultural concept of illness: Variation and validation. *American Anthropologist 86*, 341–351.