

In Search of Communities of Belief: Identifying Consensus and Disagreement in Data

Shane T Mueller, Ph.D

Applied Research Associates

and

Michigan Technological University

shanem@mtu.edu

This research has been supported in part by the Air Force Research Laboratory.

Cultural Consensus Theory

- Can tell you:
 - Whether there is agreement
 - What the agreement is
 - How “competent” individuals are
 - Identify the answer “without the answer key”.
- It is not without limitations

Limitations

- What happens when you don't find a consensus?
- What if there is more than 1 group?
 - Political factions, cohorts
- Filters everything through a correlation matrix
 - It throws away a lot of data
 - Relies on ad hoc methods to compute a correlation matrix
- Rule-of-thumb notions of convergence

Alternative: Finite Mixture Modeling

- ➔ To identify cultures of shared belief, I've adapted statistical methods for model-based clustering to develop ***Cultural Mixture Modeling***.
 - A statistical modeling approach that embeds theories of cultural transmission and agreement.
 - Model data as a mixture of distinct groups.
 - Identify how many groups there are.

Basic Procedure of CMM

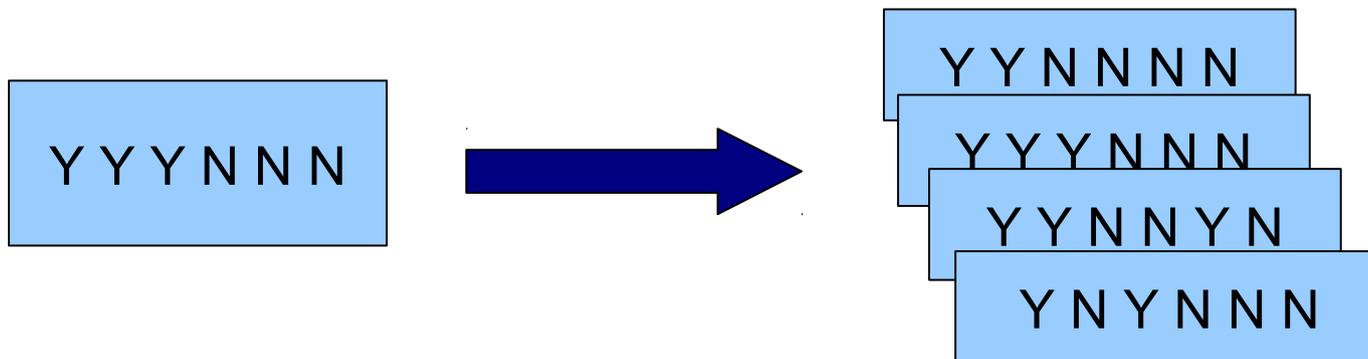
- ➔ Identify a probability model (generative model) of data
- ➔ Select fixed number of latent groups
- ➔ Use E-M optimization to identify best group members
- ➔ Repeat for 1,2,3,4,...N latent groups
- ➔ Use BIC statistic to select simplest most-descriptive model

Information Transmission: The Strong Consensus Model

- ➔ Groups and individuals have consistent set of yes-no beliefs.

Group Belief

Individual Beliefs

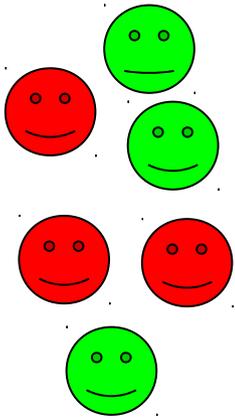


Each belief transmitted to individual with error rate γ

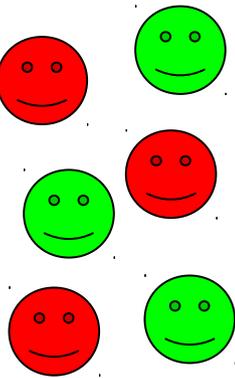
E-M Procedure: 2 Groups

Round 1

Sort Randomly

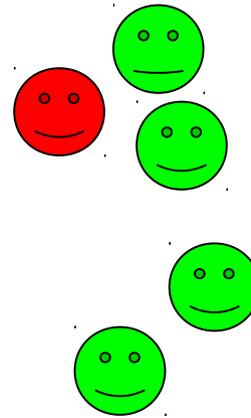


Compute Expected Belief

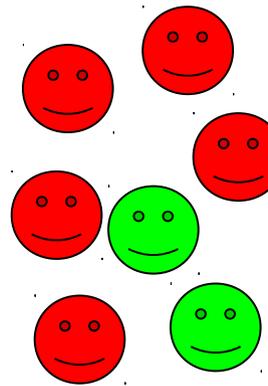


Round 2

Sort to Maximize

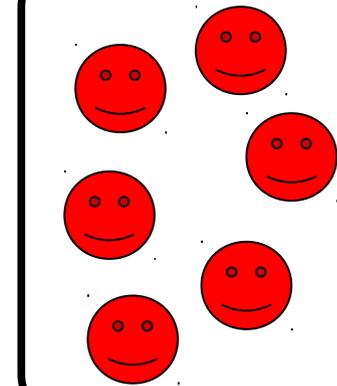
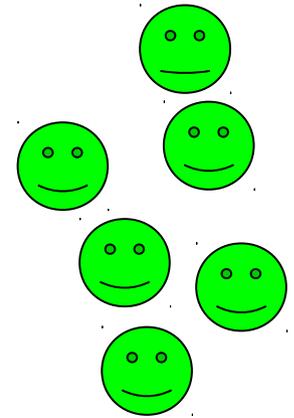


Compute Expected Belief



Round 3

Sort to Maximize



Implementing CMM

- Leisch's & Gruen's flexmix package for R handles the basics of most of this.

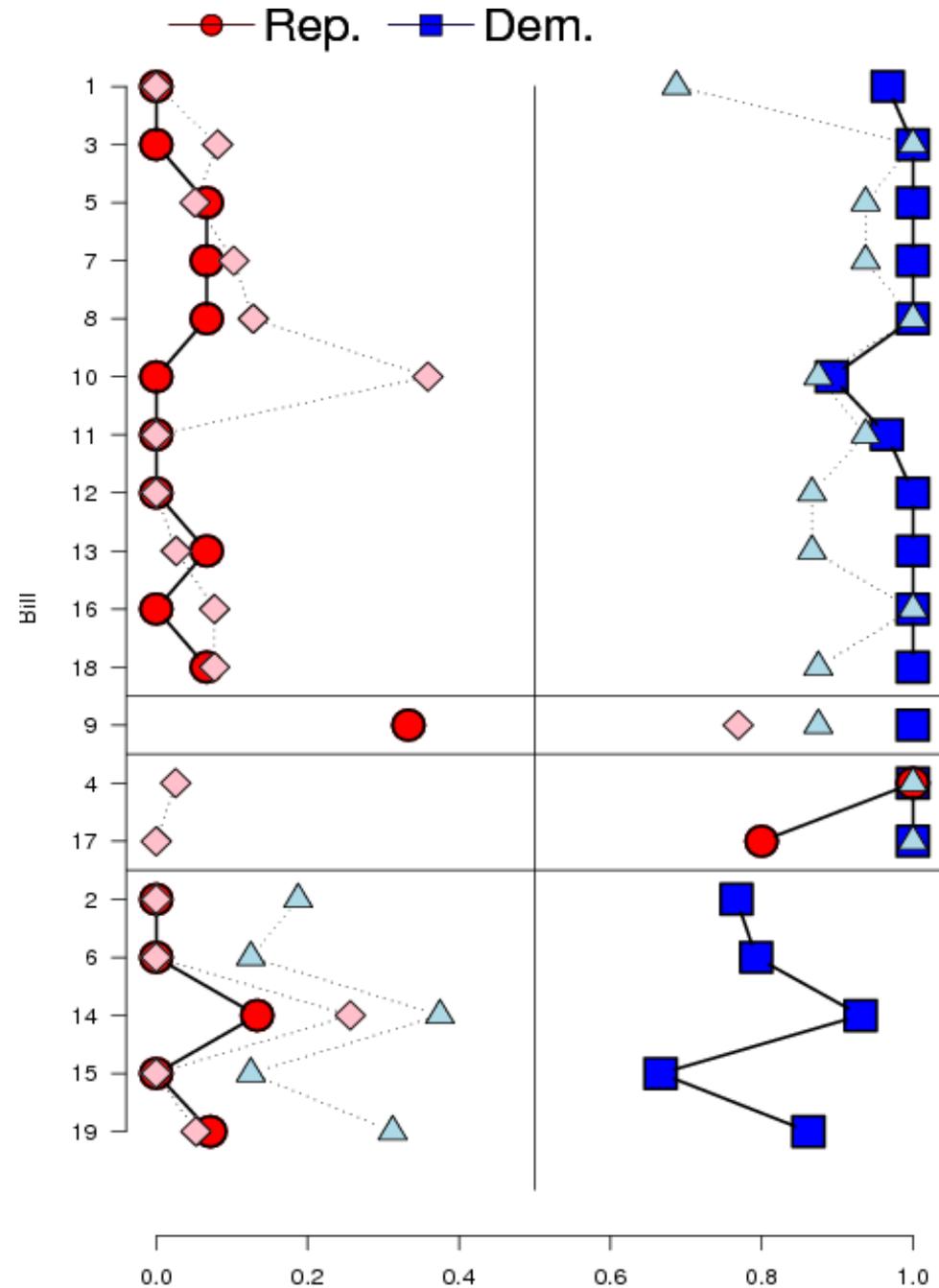
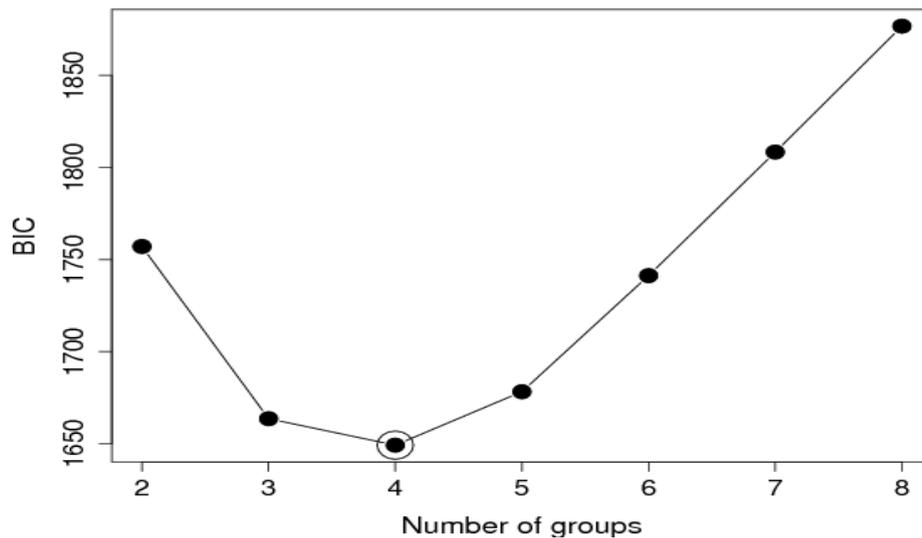
```
group <- stepFlexmix(as.matrix(dat.sub)~1,  
                    model=mymclustConsensus(),  
                    control=list(minprior=.005,classify="random"),  
                    k=1:8,nrep=15)
```

```
g1 <- getModel(group,"BIC")  
post1 <- (posterior(g1))  
c1 <- clusters(g1)  
tab1 <- table(c1)  
ps1 <- parameters(g1)  
ctab <- cumsum(tab1)
```

Applications

2006 US Senate votes on Labor

- ➔ 4 groups: 2 Rep, 2 Dem
- ➔ Each party consists of larger more liberal group & smaller more conservative group
- ➔ Cultural labels (party) fail to capture true structure



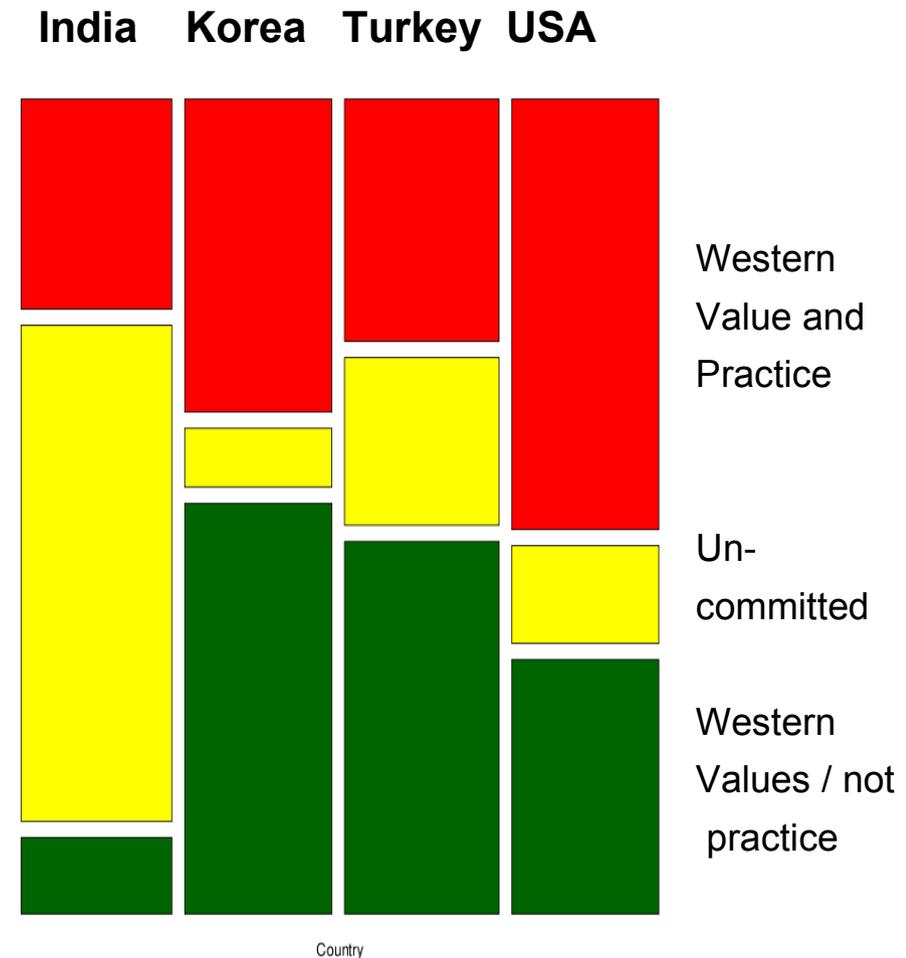
Beliefs about Teamwork

➔ Survey of business leaders in 4 cultures (US, Turkey, Korea, India)

➔ 66 Questions about teamwork.

“To make a decision, the team discusses and debates different ideas and votes to make the final decision.” (Both prescriptive and descriptive judgments)

➔ Results: belief groups crossed nationalities.



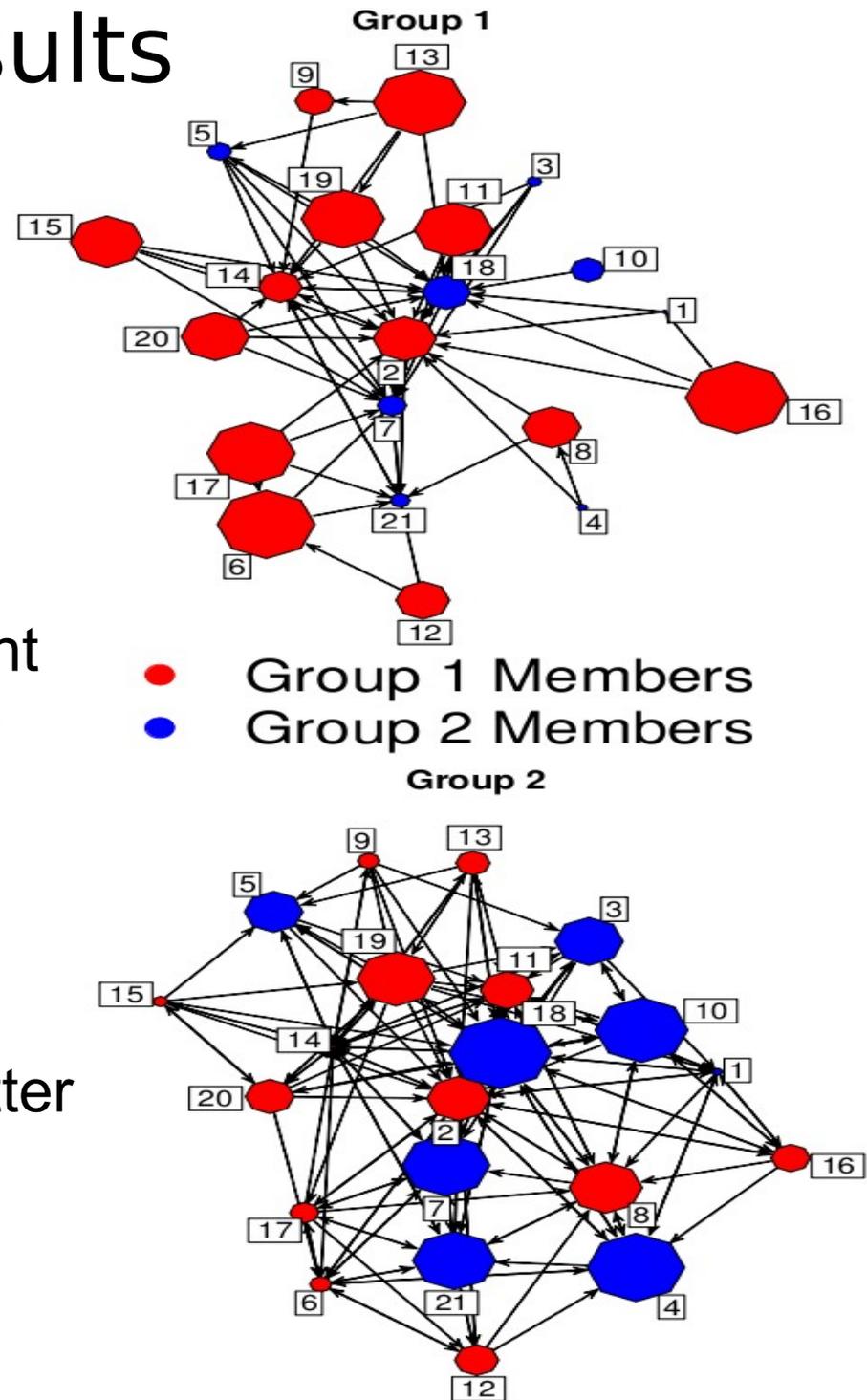
Krackhardt's (1987) Cognitive Social Structures

- Can you apply this to more complex data?
- Krackhardt studied a social advice network among 21 managers.
- He collected what he called a “Cognitive Social Structure” a complete social network from each person.
- Result was 21 21x21 networks.

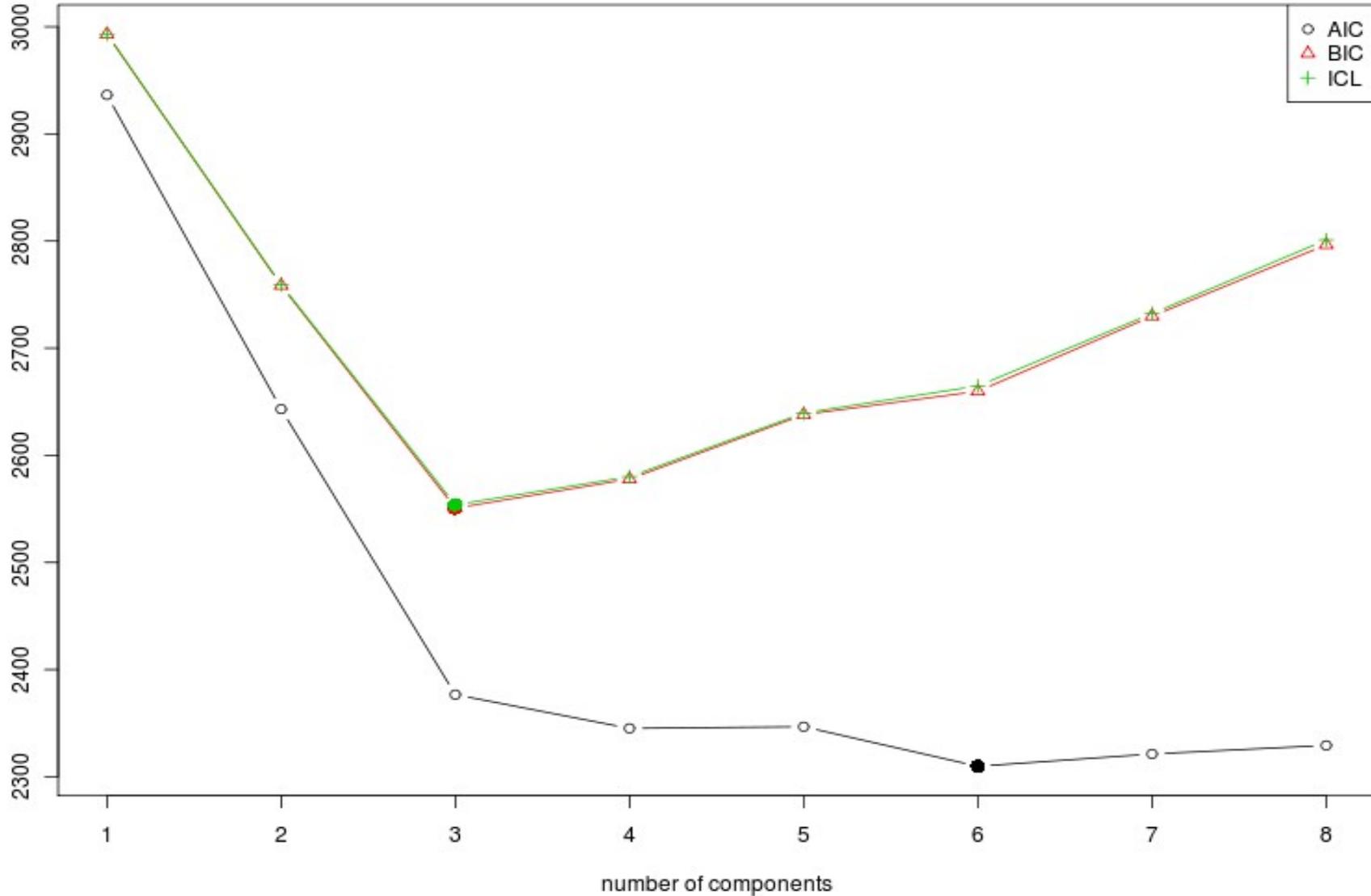
- Applying CMM involves treating each network as 21² 'questions'
- He claimed that there was a consensus and so collapsing them was appropriate.

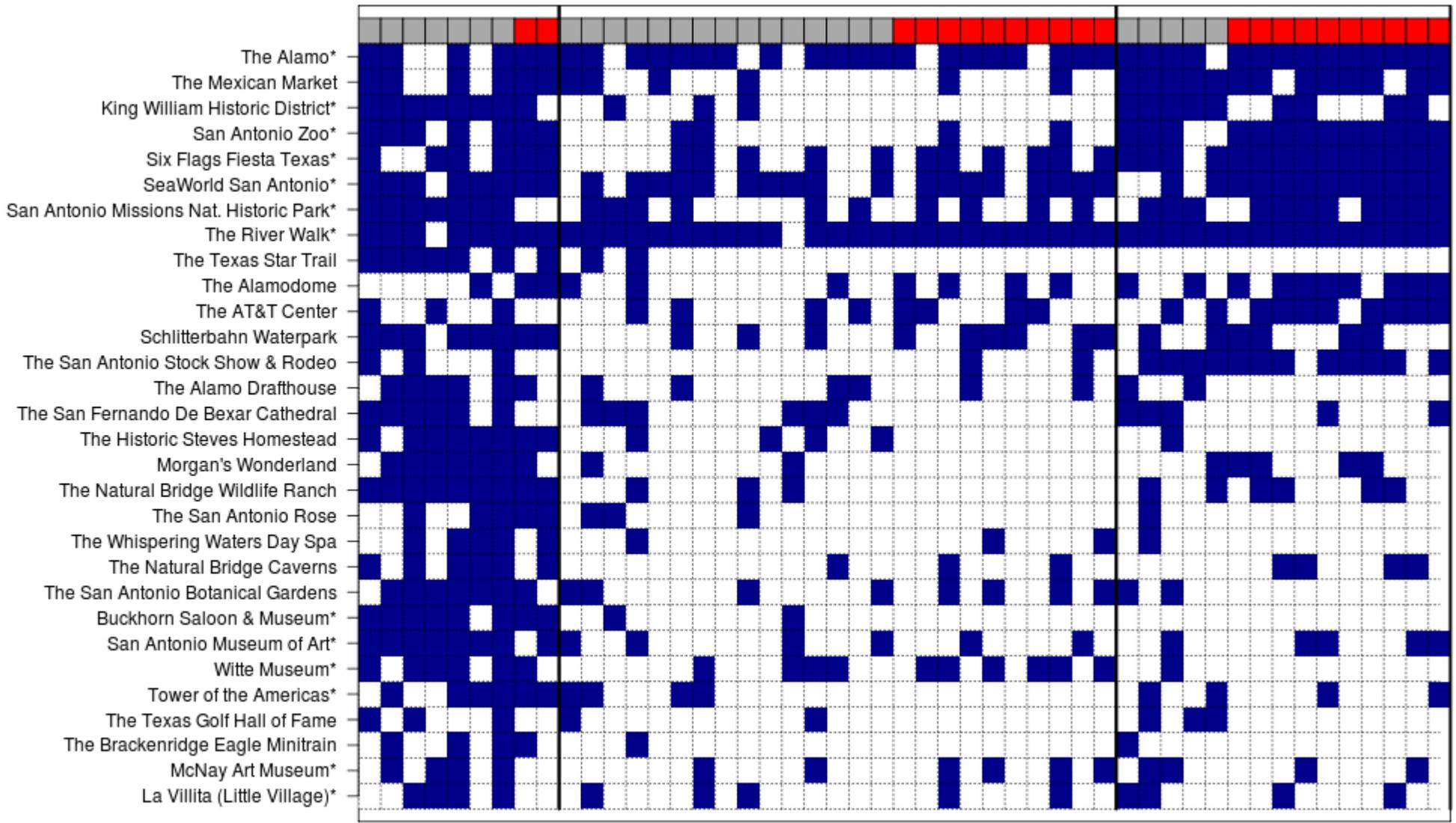
Krackhardt (1987) Results

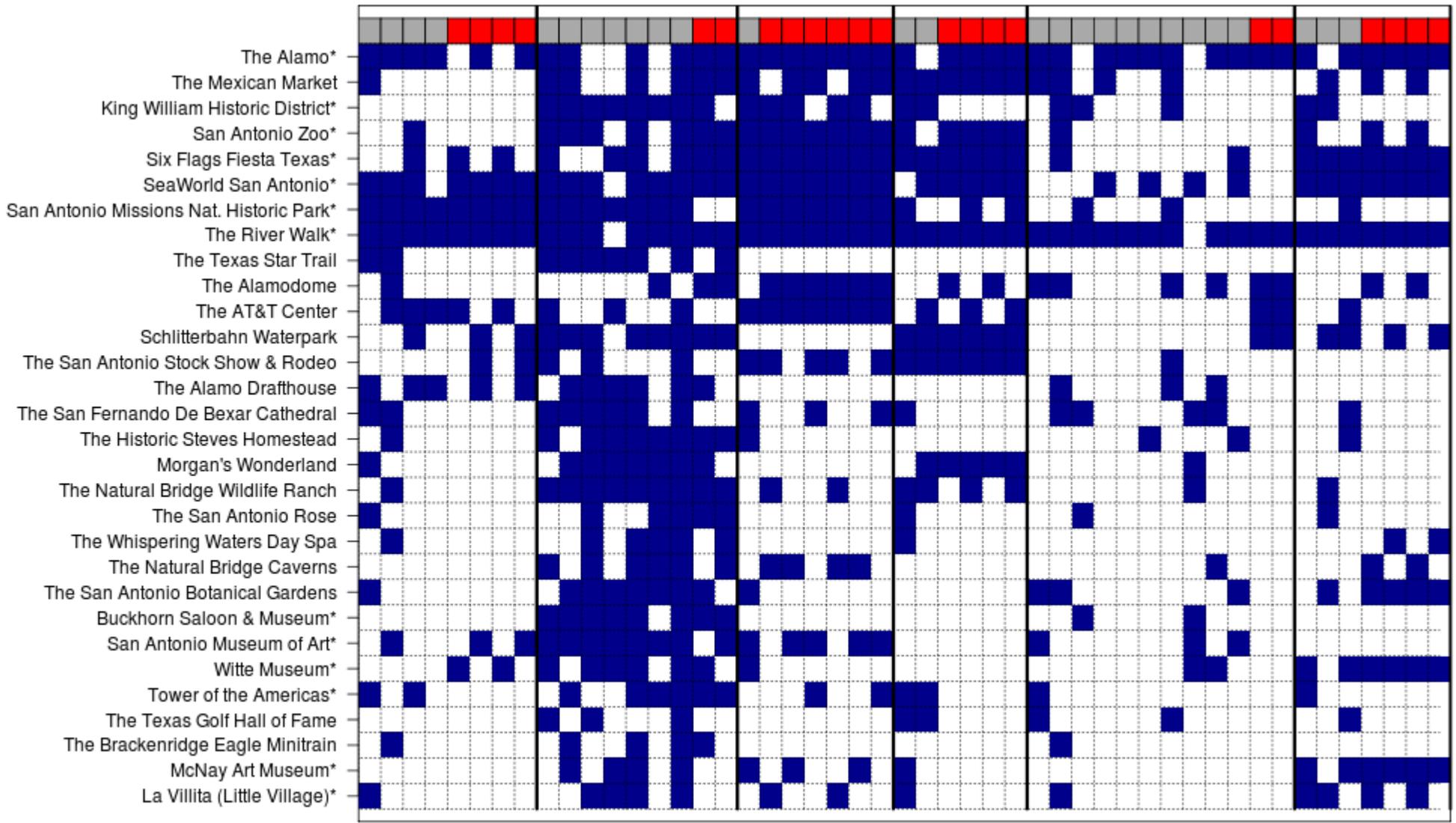
- Two networks emerged: “Democratic” and “Hierarchical”
- The more central managers believed in a more democratic advice network.
- The peripheral managers thought that advice was reserved for the central managers.
- The “ground truth” (locally-aggregated) network suggested that trust was democratic.
- The central managers had a better understanding of the complete network.



San Antonio Survey







Summary

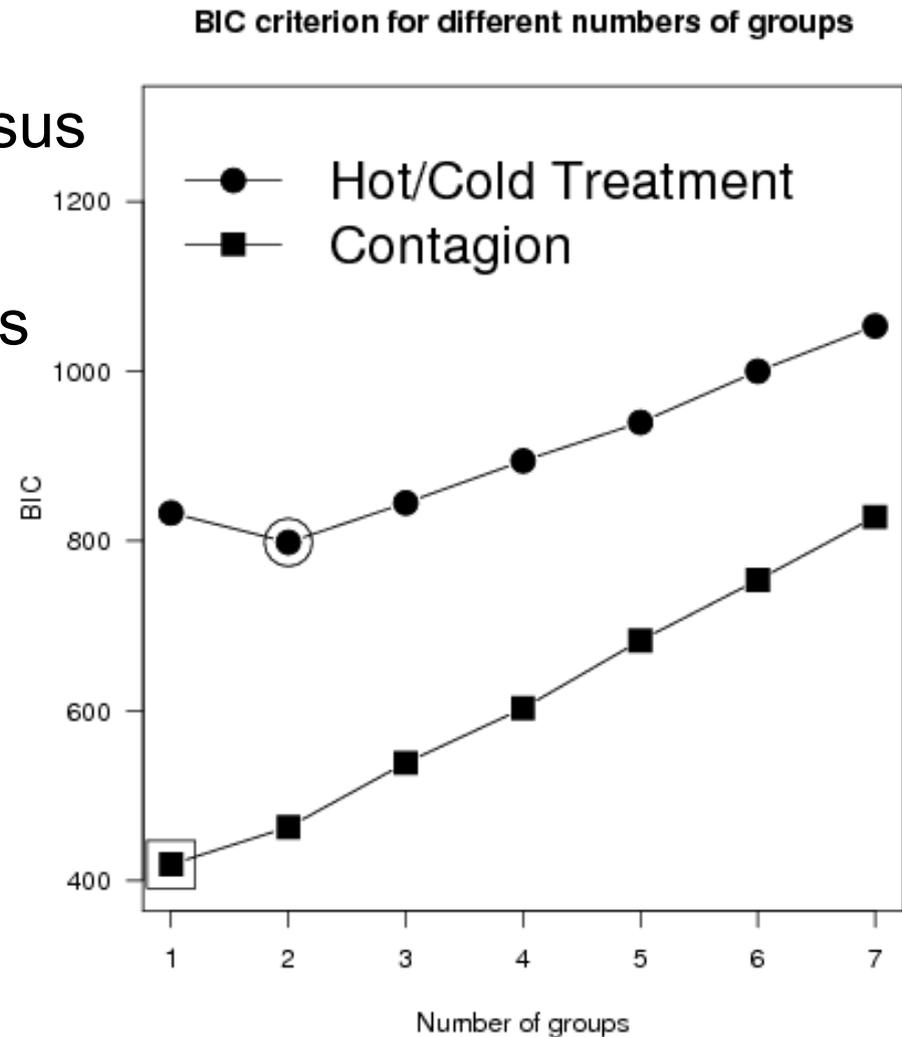
- “Cultural Mixture Modeling” is a version of CCT that aims to find groups of agreement within a population.
- Allows more coherent and flexible models, unambiguous decision rules, multiple groups.
- Cost is the increased complexity of applying the models
 - R package flexmix for doing E-M optimization for finite mixture models.
 - Many bayesian mixture models allow alternative inference method.
- Using CMM offers window into “groups of belief”, as an alternative to demographic slicing-and-dicing.

Romney, et al. (1999) data from Weller (1983, 1984)

- ➔ Native Guatemalan women were interviewed about their beliefs on the causes and treatments of diseases
 - ➔ Contagious or non-contagious
 - ➔ Treated with hot or cold treatments
- ➔ Romney determined that there was consensus about contagion, but not about treatment

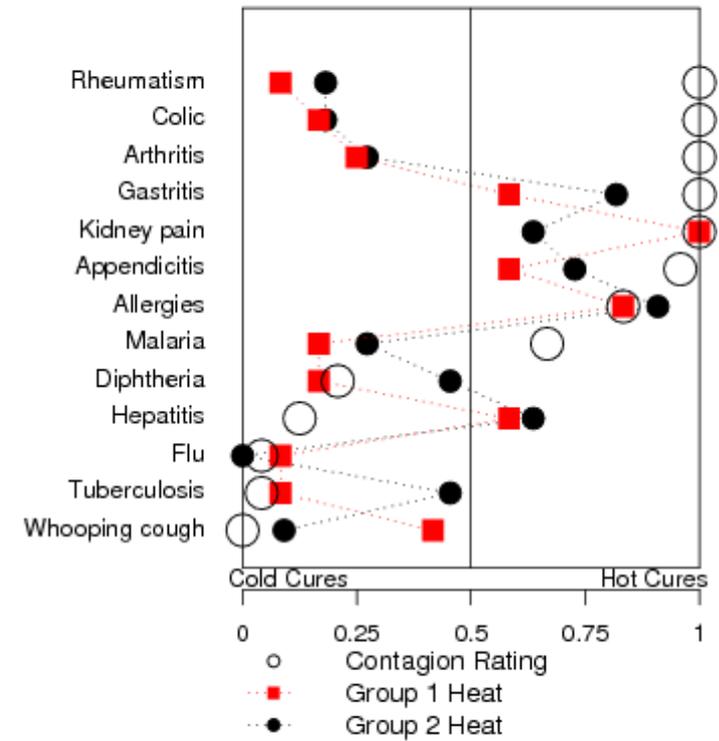
Results

- ➔ Like Romney et al. (1999), Contagion produced consensus
- ➔ Like them, treatment did not
- ➔ Treatment produced 2 groups



Inferred mental models

- ➔ There is a subset of diseases most agree on treatment
- ➔ For other diseases, respondents either:
 - ➔ Contagion \leftrightarrow heat Contagion \leftrightarrow cold
- ➔ Group membership was correlated with age & # of children



Disagreement on treatment

