

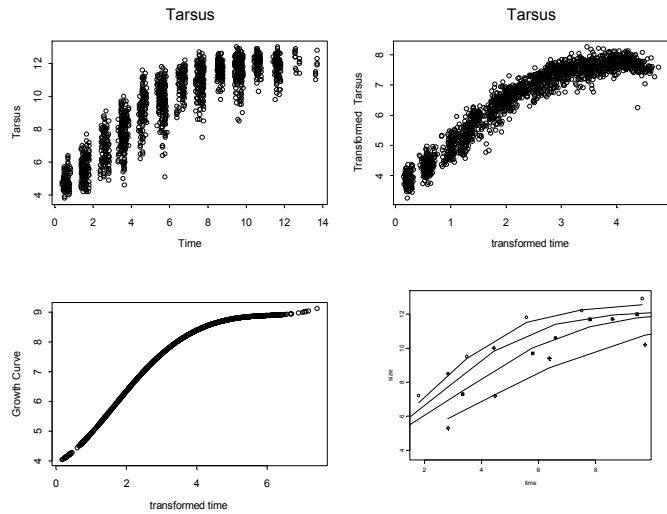
# Statistical Research

Prof. Naomi Altman

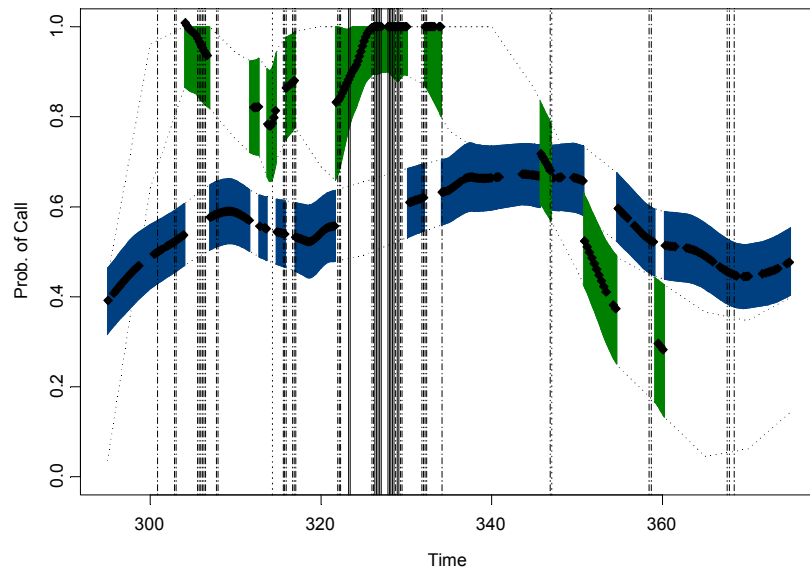
Undergraduate Statistics Club

March 19, 2008

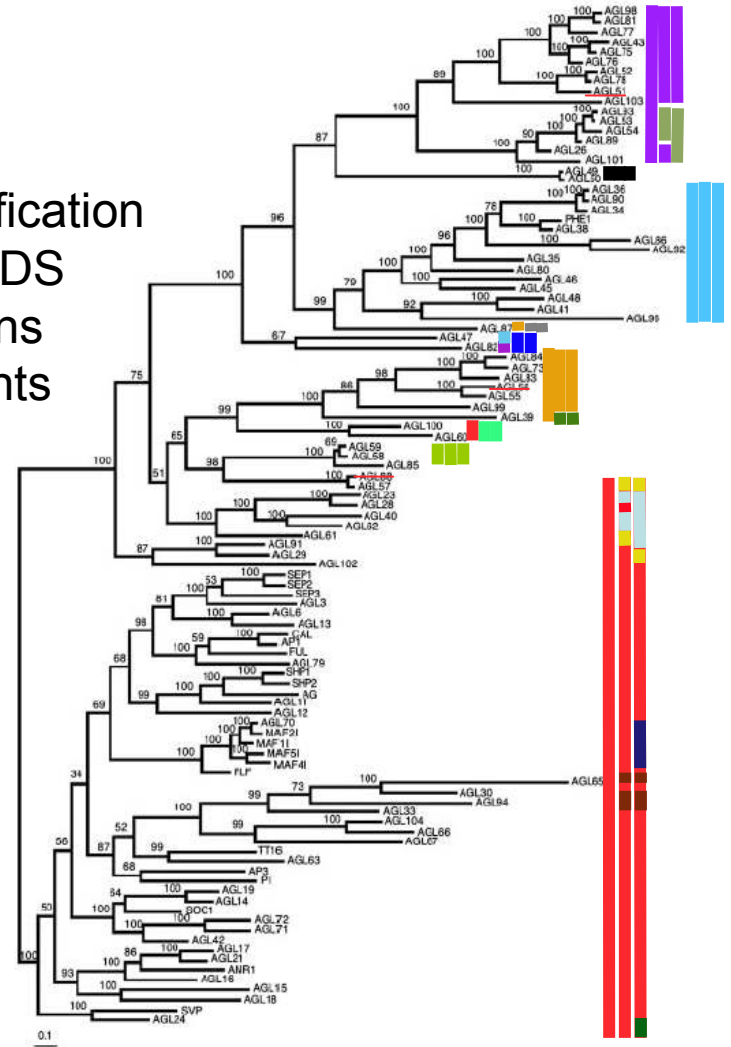
# All we need are data ...



growth of tree  
swallow chicks



classification  
of MADS  
proteins  
in plants



number of blue  
whale calls when  
sonar is off (green) or  
on (red)

# Understanding Human Interactions

How can we understand how people behave in natural settings?

e.g. Can we observe a preschool child in some setting and use this observation to predict whether the child will exhibit "problem behavior" in school?

Can we observe an engaged couple and predict if their marriage will succeed?

It is not clear what to measure.

There are many interacting "players".

The outcome may depend on unobserved variables.

# Emotional Regulation

The research is being done in Dept. of Psychology by Prof. Pamela Cole  
graduate students: Sarah Hall, Patricia Tan

Objective: To understand how the mother/child interaction works to teach a child to regulate emotion

Mother/child pairs are brought into the lab.

A wrapped toy is placed on the table and the child is told they will get the toy at the end of the session.

Mother is given a questionnaire to fill out, and child is given a boring toy to play with.

After 8 minutes, the child is given the wrapped toy.

The same pair is measured at child's age 18, 24, 36 and 48 months.

The session is videotaped. After the session the researchers use the tape to obtain a second by second record of behaviors of mother and child such as smiling, exhibiting anger, interacting with each other, child's "strategy" for keeping busy, mother's strategy for controlling child, etc.

# Objectives

Investigators observe that at 18 and 24 months, modal behavior is distressed and disruptive. At 36 and 48 months, calm behavior is more often observed. Children forget the task between visits, and continue to want the gift.

Why does behavior change?

1. Are children less eager and frustrated? (No.)
2. Are children better able to delay frustration or control anger?
3. Do children self-distract better or become more absorbed in distracting activities?
4. Do children "use" their mothers differently?

# Example of Data

TIME	NEU	HAP	ANG	FD	ODD	D1	D2	NEU	HAP	ANG	FD	ODD	D1	D2
1	0	0	1	0	0	0	0	0	1	0	0	0	0	1
2	0	0	1	2	1	0	0	1	0	0	0	1	0	1
3	0	0	1	0	0	0	0	0	1	0	0	0	0	1
4	0	0	1	0	0	1	0	0	1	0	0	1	0	0
5	0	0	1	0	0	0	0	0	2	0	0	1	0	1
6	1	0	0	0	0	1	0	0	2	0	0	0	0	1
7	0	1	0	0	0	1	0	0	2	0	1	0	0	1
8	1	0	0	1	0	1	0	0	1	0	0	1	0	0
9	1	0	NA	0	0	1	0	1	0	0	0	0	1	0
10	1	0	0	0	0	0	0	1	0	0	0	0	0	0
11	0	0	2	2	0	0	0	1	0	0	0	0	0	1
12	1	0	0	0	0	0	0	1	0	0	0	1	0	0
13	1	0	0	0	0	1	0	0	1	0	0	0	0	3
14	1	0	0	0	0	1	0	1	0	0	0	0	0	1
15	0	1	0	0	1	0	0	1	0	0	0	0	0	1

NEU=neutral HAP=happy ANG=angry FD=focus on demand ODD=other directed distraction  
D1=self-directed transient distraction D2=self-directed focused distraction

# Some ideas of how to analyze the data

1. Focus on 1 variable at a time.

e.g. time until first event, time between events (Prof. Runze Li)  
pattern of emotional intensity over time (Prof. Damla Senturk)  
variability (lability) of emotional intensity (Prof. Naomi Altman)  
level of emotional intensity (Senturk and Altman)

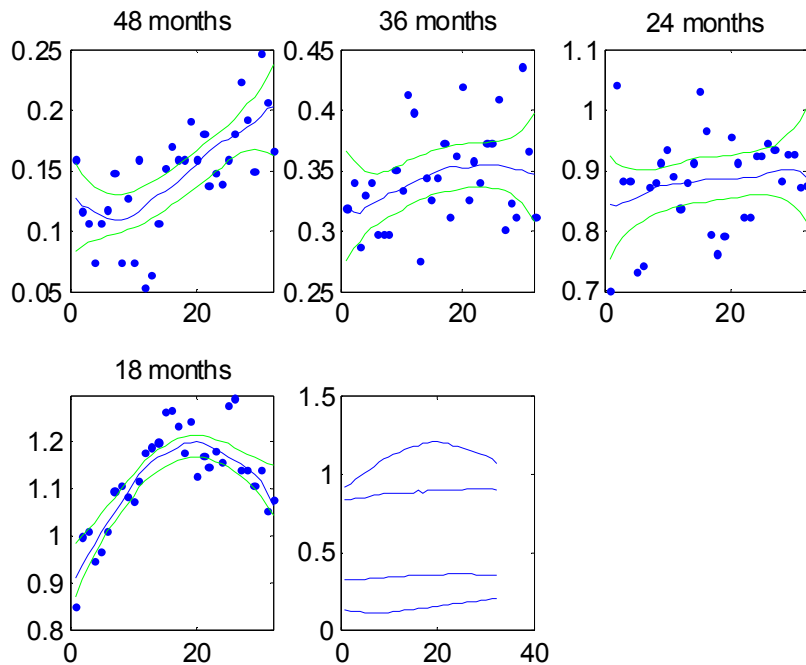
2. Combine variables.

e.g. Child cannot be happy, neutral and angry at the same time, so combine these into a single scale.

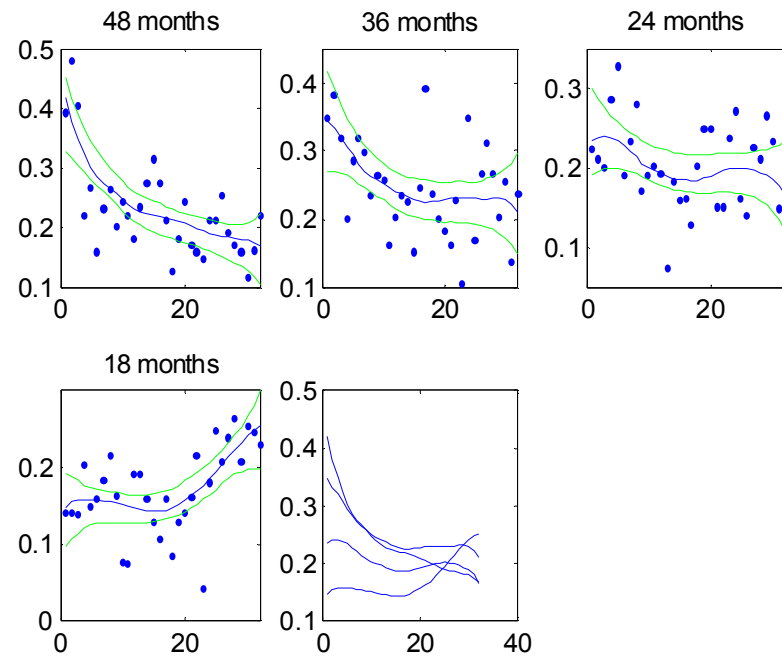
3. Look at the interaction between emotion variables (NEU, ANG, HAP) and behavior variables.

Note that we are not even looking at the mother yet.

# Emotional Intensity over Time



Average ANGER level over time at different ages



Average HAPPY level over time at different ages

Researchers thought older kids would be better able to anticipate the end of the session and would be happier towards the end.

# Nonparametric smoothing

The curves are computed by local weighted regression (loess):

Movie Clip

# But:

We want to follow individuals, not average behavior.

The actual time of an event is less important than whether or not it occurred, the intensity and how long it lasted.

Lability: the changeability of the emotional state

e.g. ANGER: 030303 is more labile than 010101  
and than 000333

On an individual basis, the important dimensions of emotion are lability and modal level.

# Measuring Lability and Modal level

time	1	2	3	4	5	6	7	8	9	10
intensity	0	0	2	1	1	0	2	2	1	0

The transition matrix is the count of the number of times the subject is in state  $i$  at time  $t$  and state  $j$  at time  $t+1$

after

before

state	0	1	2
0	1	0	2
1	1	1	0
2	0	2	1

# Measuring Lability and Modal level

time	1	2	3	4	5	6	7	8	9	10
intensity	0	0	2	1	1	0	2	2	1	0

The transition matrix is the count of the number of times the subject is in state  $i$  at time  $t$  and state  $j$  at time  $t+1$

after

state	0	1	2
0	1	0	2
1	1	1	0
2	0	2	1

before

Subjects with similar transition matrices are similar in both the time spent in each state and the lability.

We could look at clusters of individuals with similar transition matrices.

# Clustering

We want to determine if there are clusters of "similar" individuals in our sample.

The results will depend on a measure of similarity and the clustering algorithm.  
The 2 most popular methods are:

**k-means** - assumes "k" clusters, estimates cluster centers and then places each individual in the cluster with the nearest center

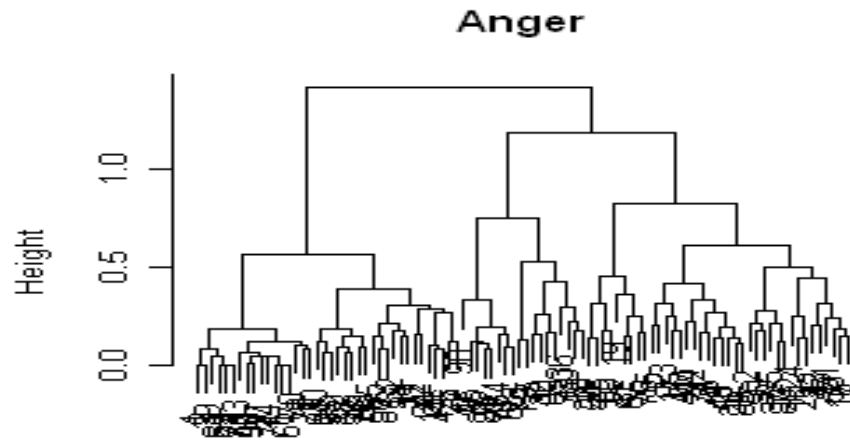
[http://home.dei.polimi.it/matteucc/Clustering/tutorial\\_html/AppletKM.html](http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/AppletKM.html)

**hierarchical agglomerative** - recursively attaches nearest neighbors to form a tree

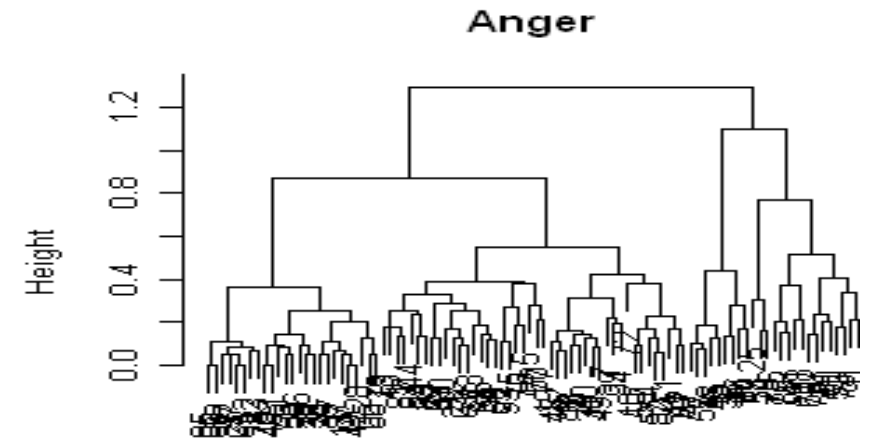
[http://home.dei.polimi.it/matteucc/Clustering/tutorial\\_html/AppletH.html](http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/AppletH.html)

Here I will use hierarchical clustering because it is more intuitive.

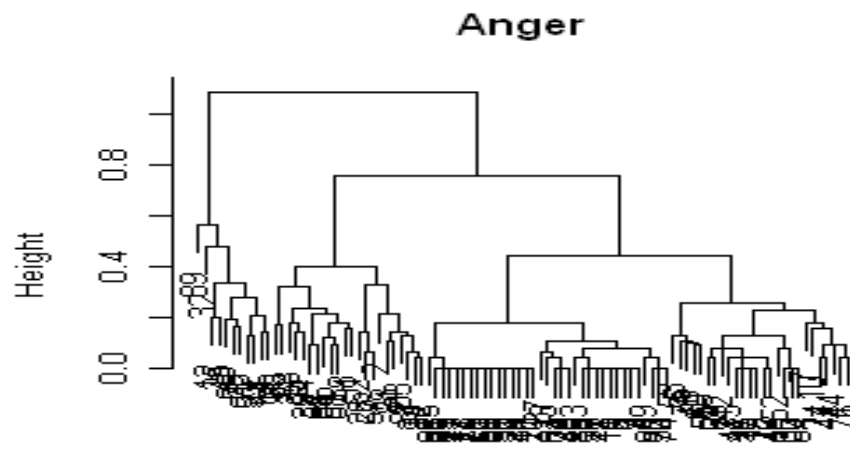
# Clustering Based on Transition Matrix



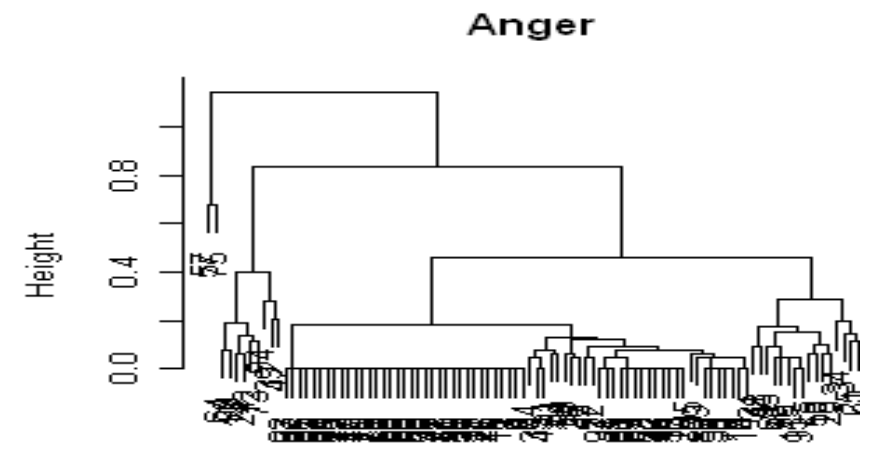
18 Mon  
hclust (\*, "complete")



24 Mon  
hclust (\*, "complete")



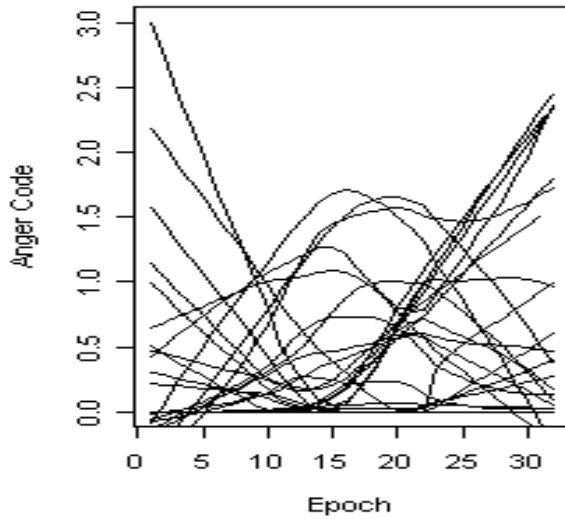
36 Mon  
hclust (\*, "complete")



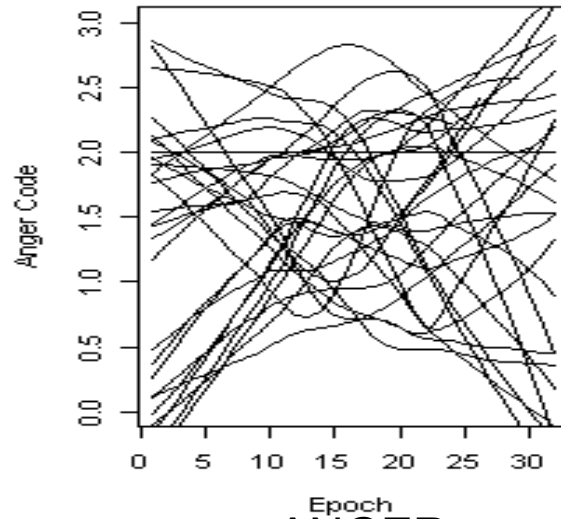
48 Mon  
hclust (\*, "complete")

# Understanding the Clusters

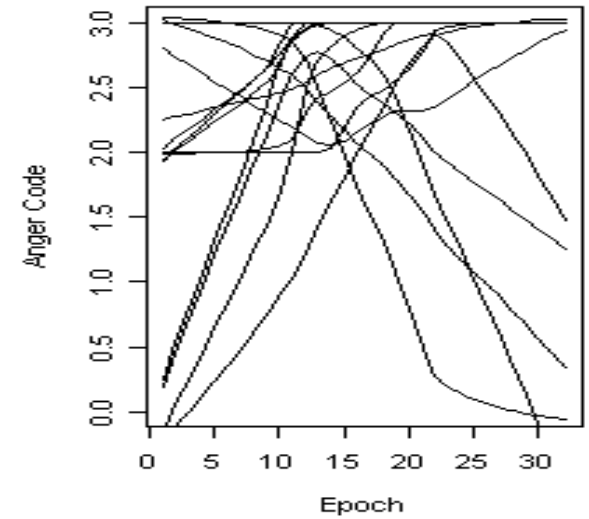
Cluster 1 Age 18



Cluster 2 Age 18

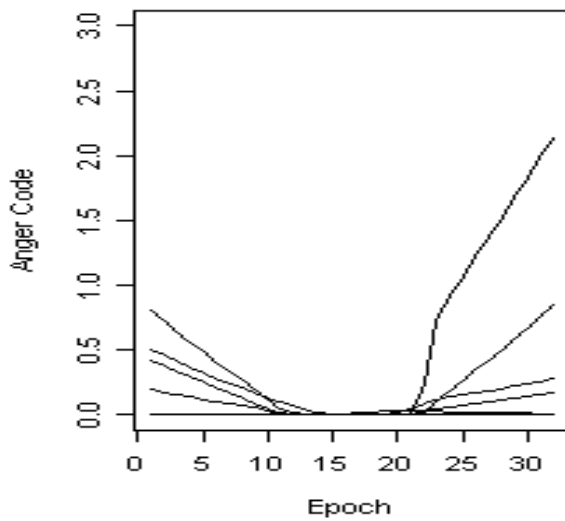


Cluster 3 Age 18

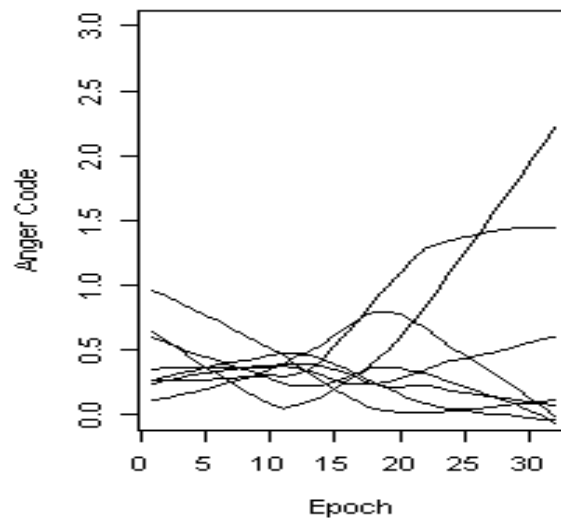


Epoch  
ANGER

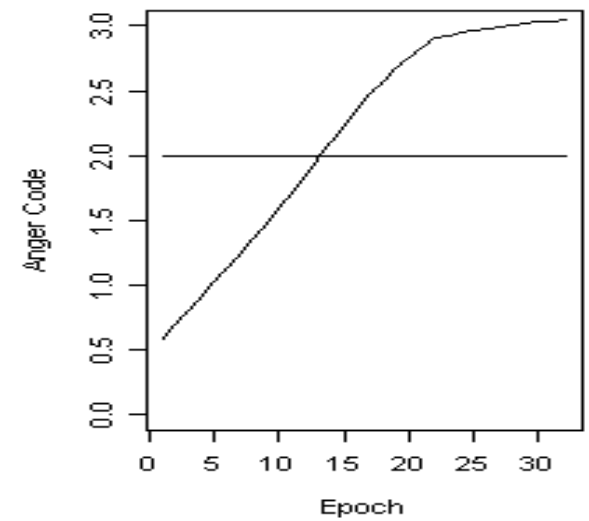
Cluster 1 Age 48



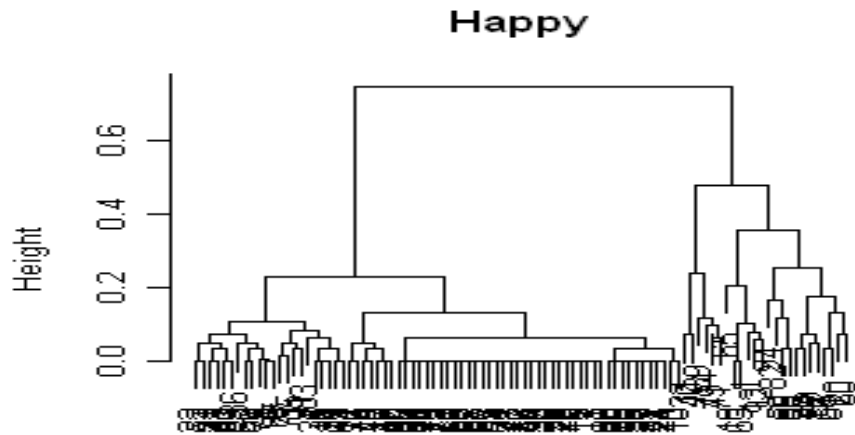
Cluster 2 Age 48



Cluster 3 Age 48



# Clusters for "Happy"



18 mon  
hclust (\*, "complete")



24 mon  
hclust (\*, "complete")



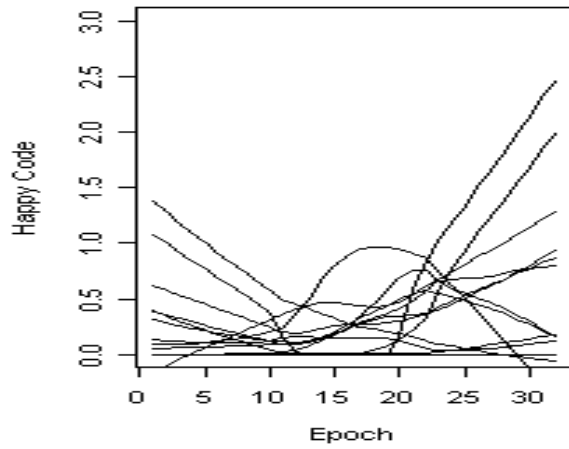
36 mon  
hclust (\*, "complete")



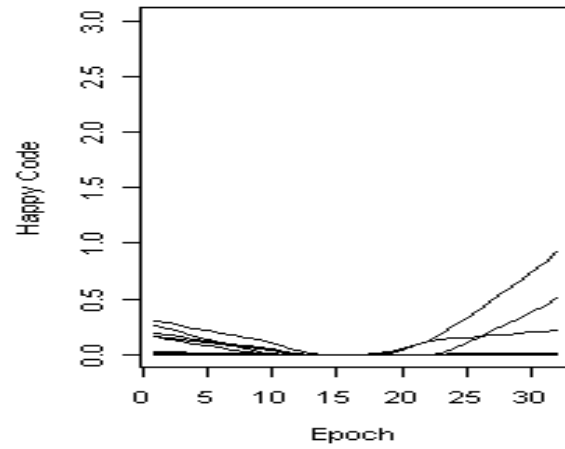
48 mon  
hclust (\*, "complete")

# Clusters for "Happy"

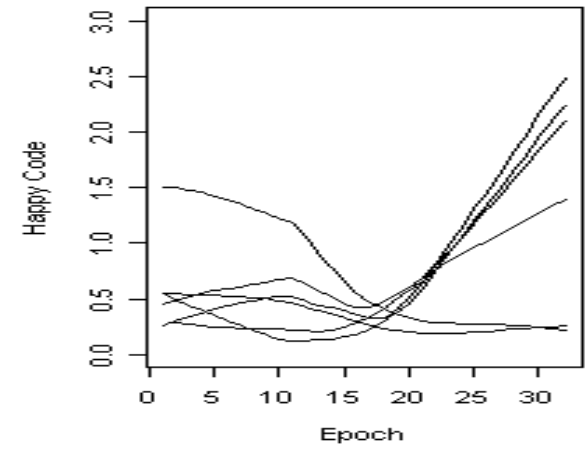
**Cluster 1 Age 18**



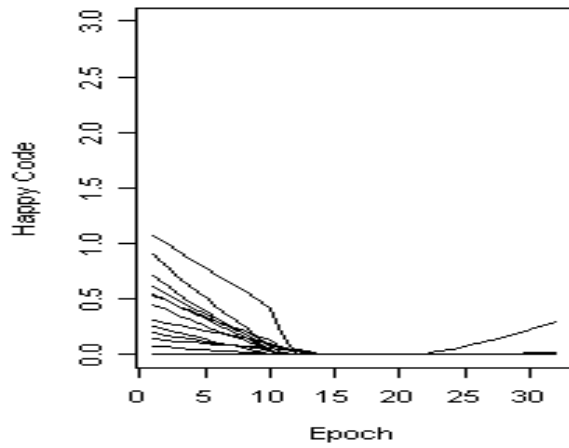
**Cluster 2 Age 18**



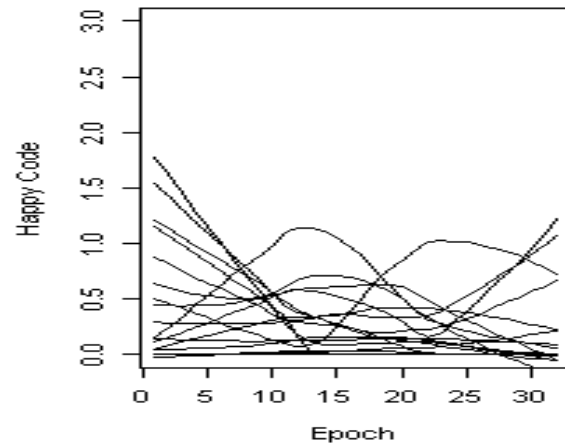
**Cluster 3 Age 18**



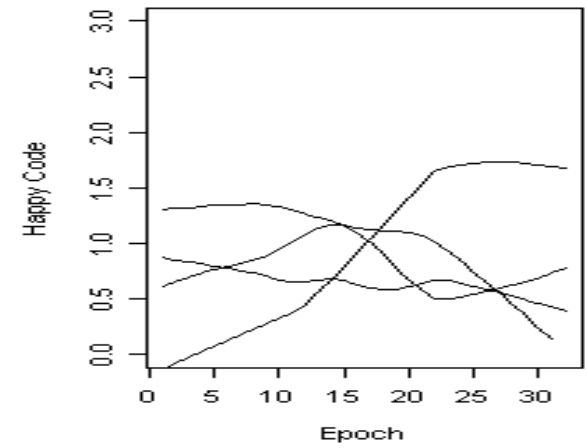
**Cluster 1 Age 48**



**Cluster 2 Age 48**

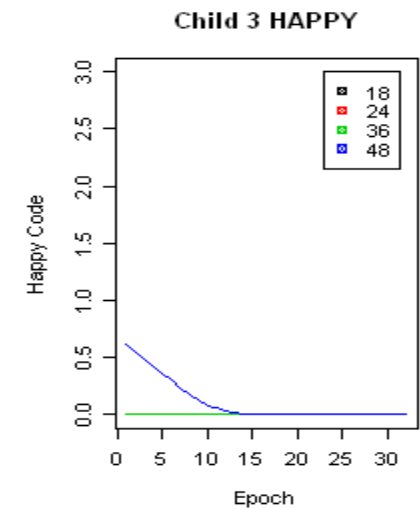
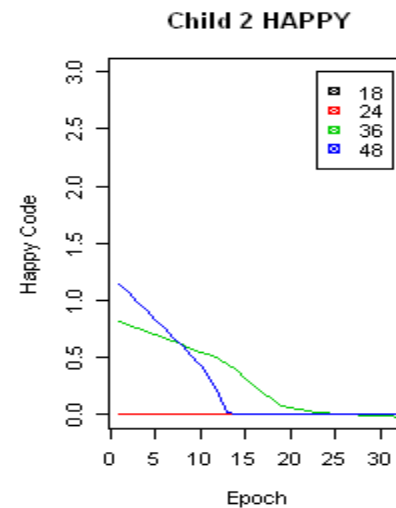
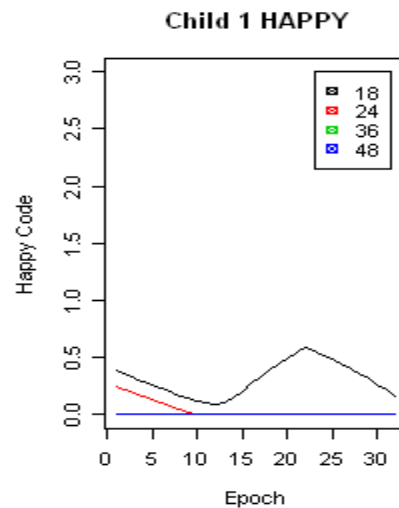
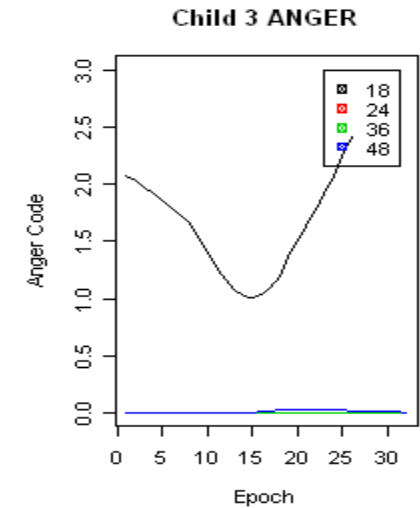
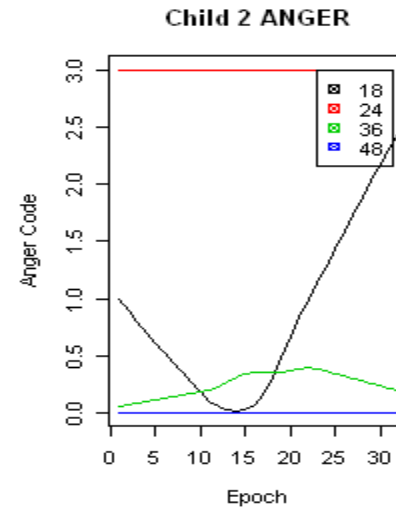
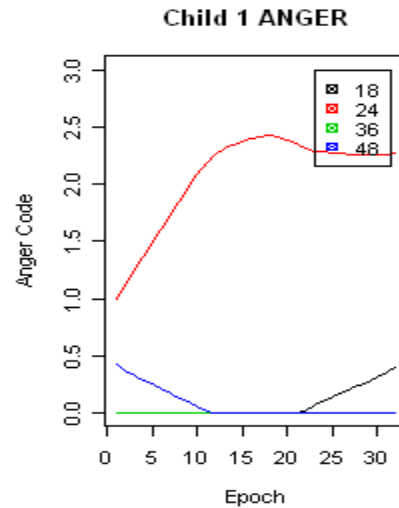


**Cluster 3 Age 48**



# Some things to do next

How consistent is the same child at different ages?  
(Not very)



# Some things to do next

Do the behavior variables predict the emotion variables (or vice versa?)

Add mom's behavior (not yet coded).

These will require statistical models such as linear or nonlinear regression.

# You can learn more

- Thursday, March 27, 2008 - 4:00 pm, 201 Thomas Bldg.
- Coffee - 3:30 pm, - 330 Thomas Bldg.

MAKRAM TALIH - Hunter College

## "STATISTICAL MODELS FOR THE SEQUENTIAL DYADIC INTERACTION IN JAPANESE AND US MOTHER-PRESCHOOLER DYADS"

Interactions between mothers and their preschool age children provide insight into how a child develops to become both a unique individual and a competent member of society. However, cultural variations in self and society may dictate distinct socialization procedures. For instance, U.S. society generally emphasizes the self as "autonomous" -- characterized by notions of competition, dominance, self-actualization and uniqueness. In contrast, Japanese society generally emphasizes the self as "related" -- characterized by notions of cooperation, empathy, harmony and accommodation.