

# Avatars and live videoconference Manipulation

Timothy R. Brick<sup>1</sup> & Steven M. Boker<sup>2</sup>

<sup>1</sup>Institute for CyberScience, QuantDev, & SSRI  
Human Dynamics and Family Studies  
Penn State University

<sup>2</sup> Human Dynamics Laboratory  
Department of Psychology  
University of Virginia

Workshop on Computationally Intensive Modeling  
of Social Interaction  
November 7, 2014



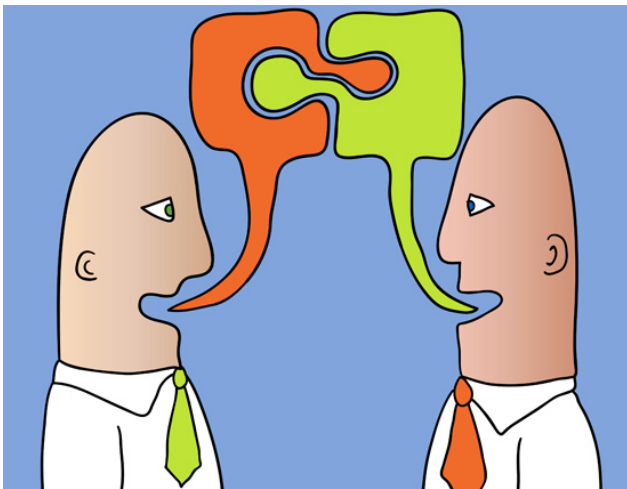
# A List of Thanks

to my collaborators and colleagues

- ▶ Allison Gray and the UVa Undergrad Team (UVa)
- ▶ Roland Göeke and his students (U Canberra)
- ▶ Jeffrey R. Spies (Center for Open Science)
- ▶ Andreas Brandmaier (MPI-Bildungsforschung)
- ▶ Angela Staples (Indiana U)
- ▶ Michael D. Hunter (U. of Oklahoma)
- ▶ Barry-John Theobald (University of East Anglia)
- ▶ I. Dzobiek, N. Green (Humboldt Univ. Berlin)
- ▶ J. Cohn (CMU), Iain Matthews (Disney)



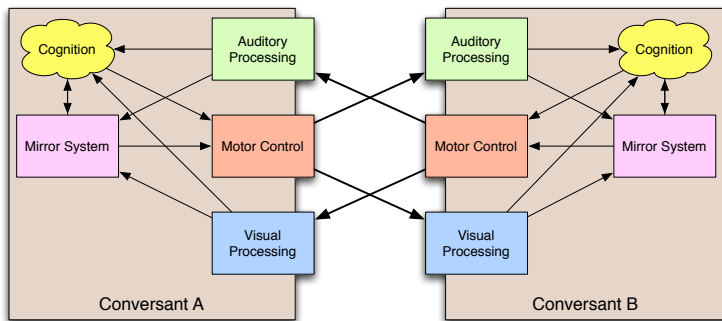
# Conversation



(Image courtesy of [www.businessgrow.com](http://www.businessgrow.com))

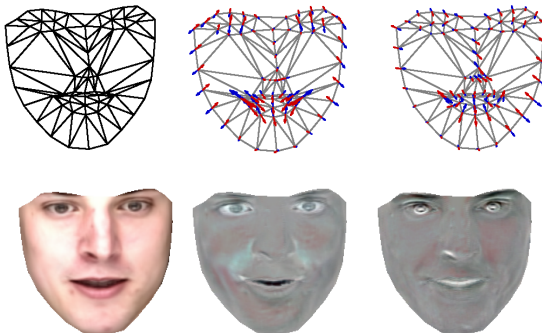
# The Model

What can we do with this model?



# How Does It Work?

## Active Appearance Models



$$\tilde{\mathbf{s}} = \mathbf{s}_0 + \mathbf{s}_s \mathbf{p}_s$$

$$\tilde{\mathbf{a}} = \mathbf{a}_0 + \mathbf{a}_a \mathbf{p}_a$$

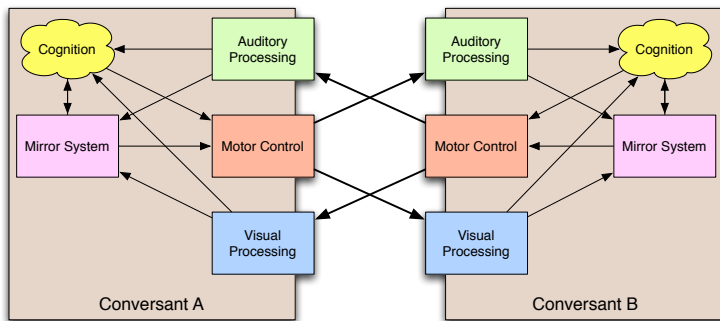
(Cootes et al., 2001)





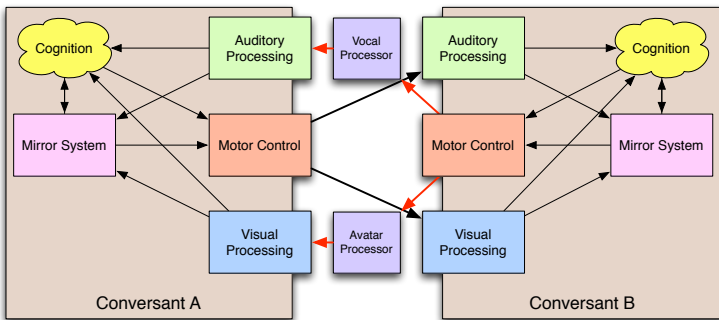
# The Model

What can we do with this model?



# The Model

What can we do with this model?







# Avatar Intervention #2:

What did we see?

- ▶ Mixed-model regression, grouped by participant, show:
  - ▶ Men move less
  - ▶ Everyone moves less if talking to a man (but not to a male avatar)
  - ▶ If I move less, you move more (Over any minute span)
  - ▶ **What matters to how you move is not who you THINK you're talking to, but how the other person moves.**



# Avatar Intervention #2: Dynamic Manipulation

- ▶ So we changed static identity
- ▶ Can we modify dynamics?
  - ▶ Can we modify dynamics?
  - ▶ Well, Kinda.



(Boker et al., 2009)



# Avatar Intervention #3: Dynamic Manipulation

## Perturbing the system

- ▶ Using RMS vertical and horizontal head velocity (captured by motion-tracking)
- ▶ Mixed-model regression, grouped by participant, show:
  - ▶ If I move less, you move more (Over any minute span)
  - ▶ But if one of us is damped, we BOTH move more
  
- ▶ What about changing between emotions?

(Boker, et. al., 2009b))



# Affective Spaces

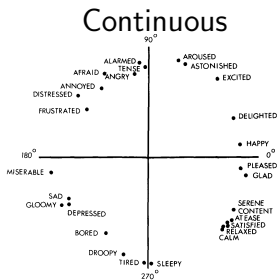


Figure 2. Direct circular scaling coordinates for 28 affect words.

**Emotion Dimensions**  
(e.g., Russell, 1980; Fontaine, et. al., 2007)

## Discrete

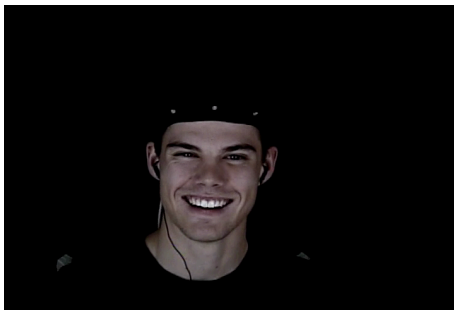


**“Big 6” Emotions**  
(e.g., Ekman, 1974; Du, Tao & Martinez, 2014)



# Affective Ratings

- ▶ 48 5-second clips from natural videoconference
- ▶ 120 participants, 30 emotion words, 9 element rating scale



## Contentment

- 
- 1 - Not appropriate at all
  - 2 - Mostly not appropriate



# Interesting Items About Emotion

1. 8-factor space
  - ▶ Joy, Anger, Shame, Disguise, Sadness, Calm, Anxiety, Interest
2. An “interest” factor
  - ▶ Surprise
  - ▶ Interest
  - ▶ Compassion



# Interesting Items About Emotion

1. 8-factor space
  - ▶ Joy, Anger, Shame, Disguise, Sadness, Calm, Anxiety, Interest
2. An “interest” factor
  - ▶ Surprise
  - ▶ Interest
  - ▶ Compassion
3. 47 out of 48 clips: significantly non-zero on more than one factor
  - ▶ 33 out of 48 clips: more than three emotions
  - ▶ Unclear whether this is concurrent or consecutive
4. Still working on the real-time mapping into this space

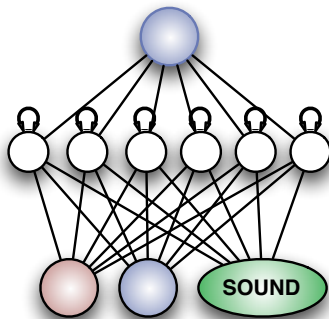


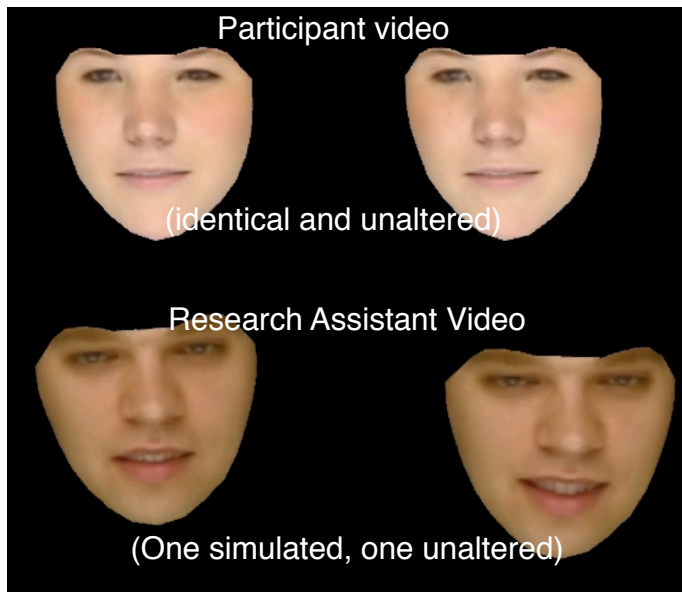


# Avatar Intervention #3: Making Faces

But not at people

- ▶ Recorded AAM Data
- ▶ Recurrent Neural Network
  1. Input: Last 8 shape parameters from both participants plus sound level in that booth
  2. Output: Shape of RA in next frame

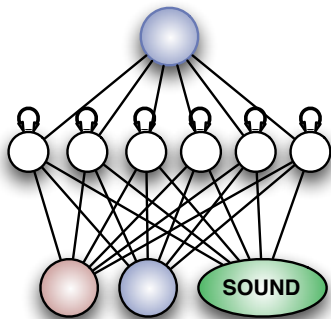




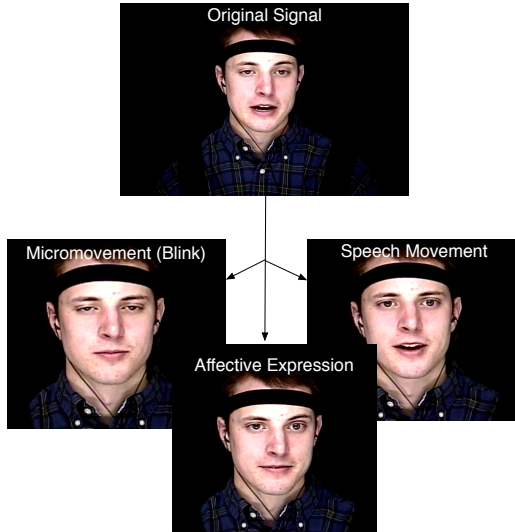
# Avatar Intervention #3: Making Faces

But not at people

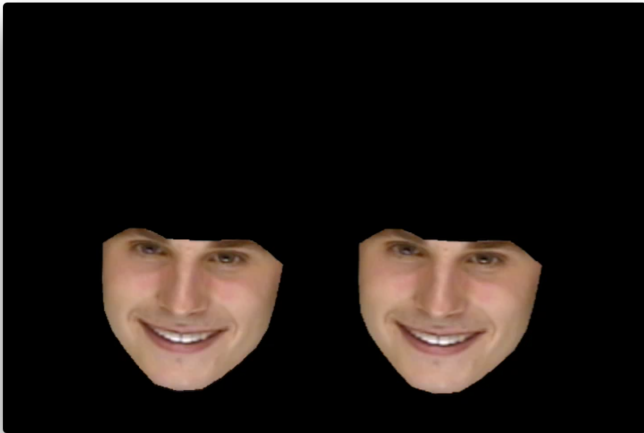
- ▶ Recorded AAM Data
- ▶ Recurrent Neural Network
  1. Input: Last 8 shape parameters from both participants plus sound level in that booth
  2. Output: Shape of RA in next frame
- ▶ Anecdotal results: about 30% fooled (without audio)



# Avatar Intervention #4: Decomposing movements



# Avatar Intervention #4: Decomposing movements



# Open Questions

- ▶ What else to do with the data
  - ▶ New analyses?
- ▶ How best to model interactive affect?
  - ▶ New generative models?
  - ▶ Time-series prediction & modification?
- ▶ How to tackle the multiscale problem?
  - ▶ Generative, interactive multilevel models?
  - ▶ How to handle speech?
- ▶ What if we want to manipulate timing?
  - ▶ Dynamic time warping?

## Thank You.

Timothy Brick <tbrick@psu.edu>



- Boker, S., Cohn, J., Theobald, B., Matthews, I., Brick, T., & Spies, J. (2009). Effects of damping head movement and facial expression in dyadic conversation using real-time facial expression tracking and synthesized avatars. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3485.
- Cootes, T., Edwards, G., & Taylor, C. (2001). Active Appearance Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 681–685.

