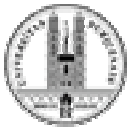


UNIVERSITY OF ZURICH

Kolloquium der Psychologie
Universität Bern, 12. Juni 2003

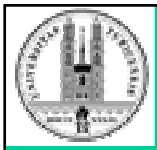
MPI FOR BIOLOGICAL CYBERNETICS

Wahrnehmung und Wiedererkennung von Gesichtern

Adrian Schwaninger

Max Planck Institute for Biological Cybernetics, Tübingen, Germany
Department of Psychology, University of Zürich, Switzerland

www.tuebingen.mpg.de
www.allgpsy.unizh.ch/Schwaninger.html



Face Recognition is very Orientation-Sensitive

Thompson (1980): „Margareth Thatcher – A New Illusion“

Rock‘s Hypothesis

Rock explains the Thatcher Illusion seven years before it has been discovered:
"In this situation [of an inverted face], there is a whole set of component figures and figural relationships to be corrected, and it is not possible to succeed in visualizing simultaneously how each of these would look were it to be egocentrically upright." (Rock, 1973, p.60)

Rock‘s Hypothesis

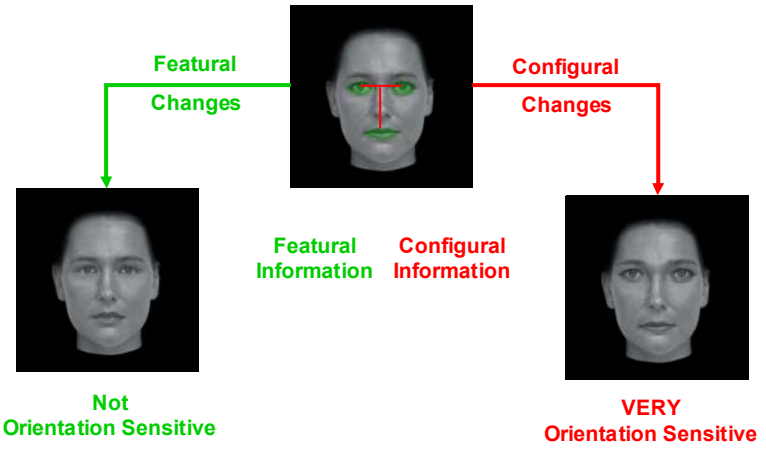
Introduction
1. Rotation
2. Parts / Conf.
3. Model
4. Computation
5. Motion
Summary

"Why is face recognition so orientation sensitive?”

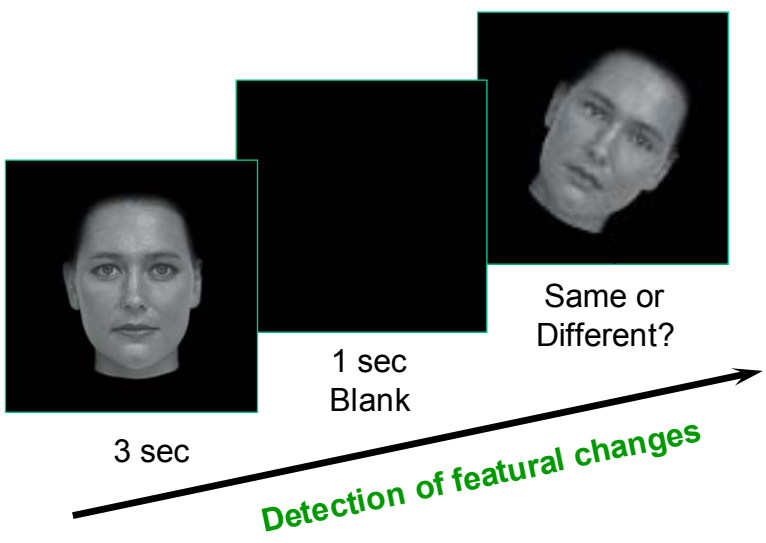
Rotated faces overtax orientation normalization mechanisms.

Rotated faces can only be processed by matching parts.

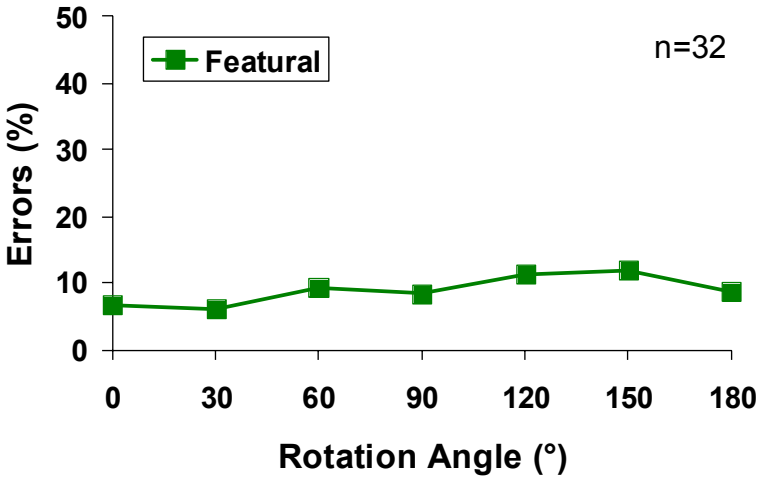
Testing Rock´s Hypothesis



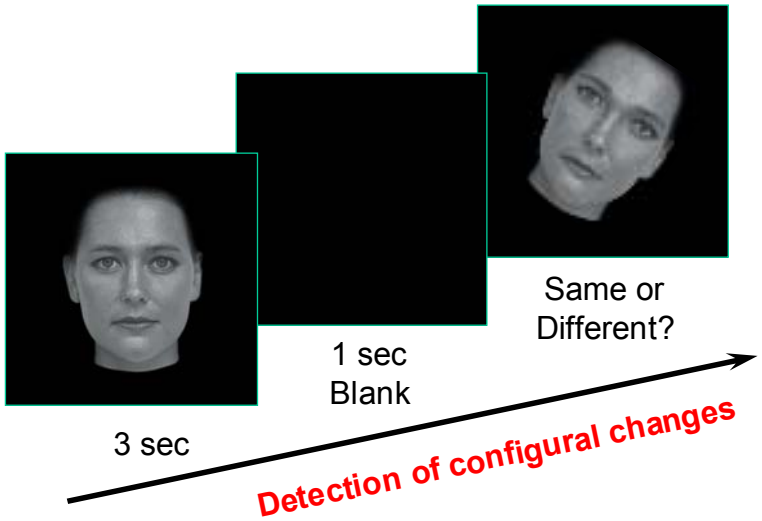
Detection of Featural Changes



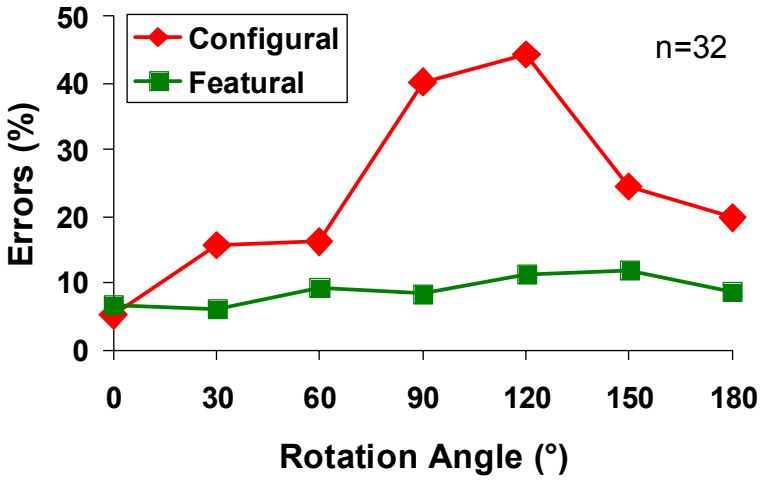
Detection of Featural Changes



Detection of Configural Changes



Detection of Configural vs. Featural Changes



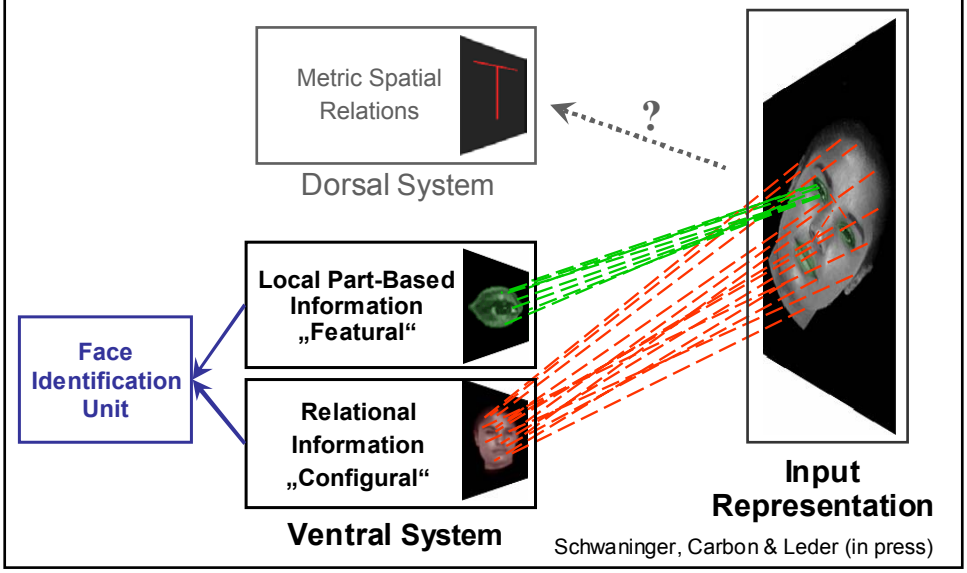
Conclusions Part 1

"Why is face recognition so orientation sensitive?"

Rotated faces overtax orientation normalization mechanisms. √

Rotated faces can only be processed by matching parts.

Integrative Model



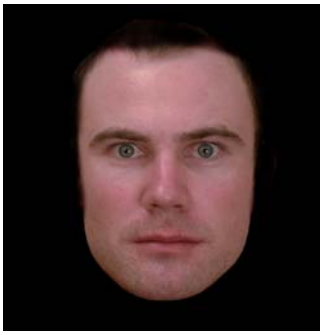
Schwaninger, Carbon & Leder (in press)

Part 2: Role of Parts And Configural Information

- Faces have often been cited as examples for **exclusive holistic** processing
  (e.g. Farah et al., 1995; Tanaka & Farah, 1991, 1993; Biederman & Kalocsai, 1997)
- In this case **holistic** means **no** explicit representations of **parts**

Parts in Face Recognition ?

Encoding Condition



10 Faces
Presented Sequentally

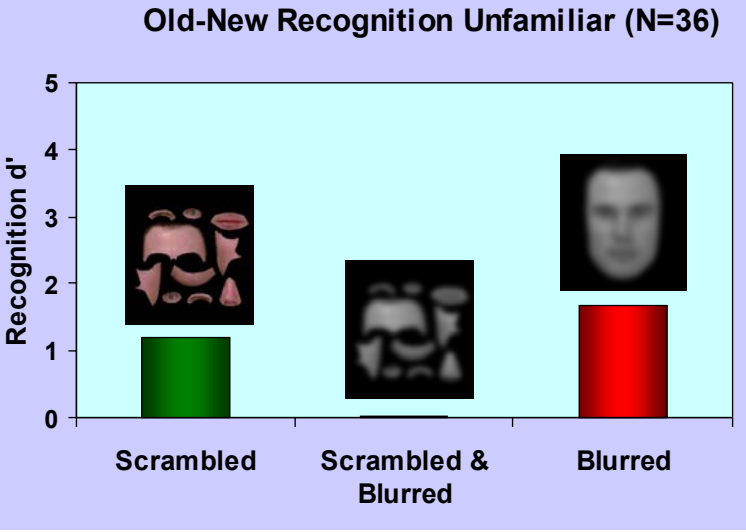
Parts in Face Recognition ?

Introduction
1. Rotation
2. Parts / Conf.
3. Model
4. Computation
5. Motion
Summary

Test Condition

10 targets („old“), 10 distractors („new“)
Decision: Old or new?

Role of Parts And Configural Information



Schwaninger, Lobmaier & Collishaw (2002)

Familiar vs. Unfamiliar Face Recognition

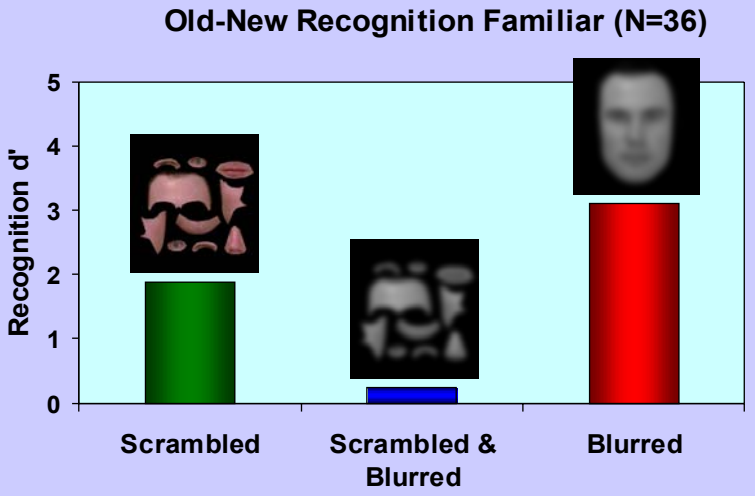
Method and Procedure

Same as Experiment 1 but participants were fellow students of the persons depicted in target faces.

→ Target faces were all familiar

→ Distractor faces were all unfamiliar

Familiar vs. Unfamiliar Face Recognition

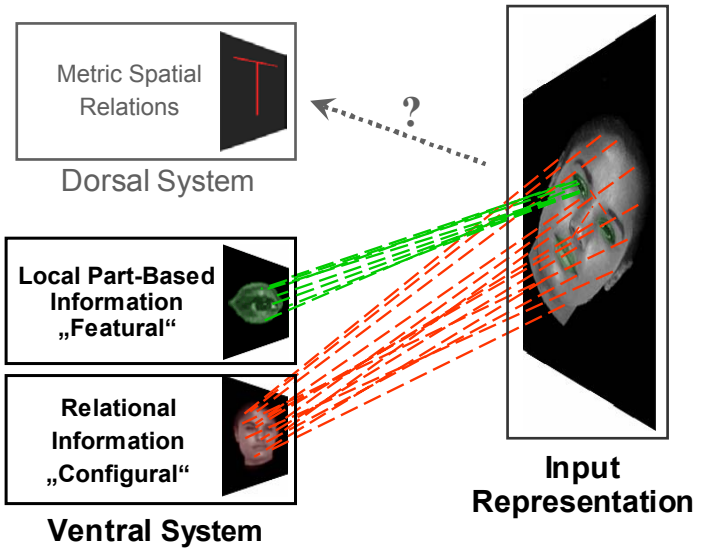


Schwaninger, Lobmaier & Collishaw (2002)

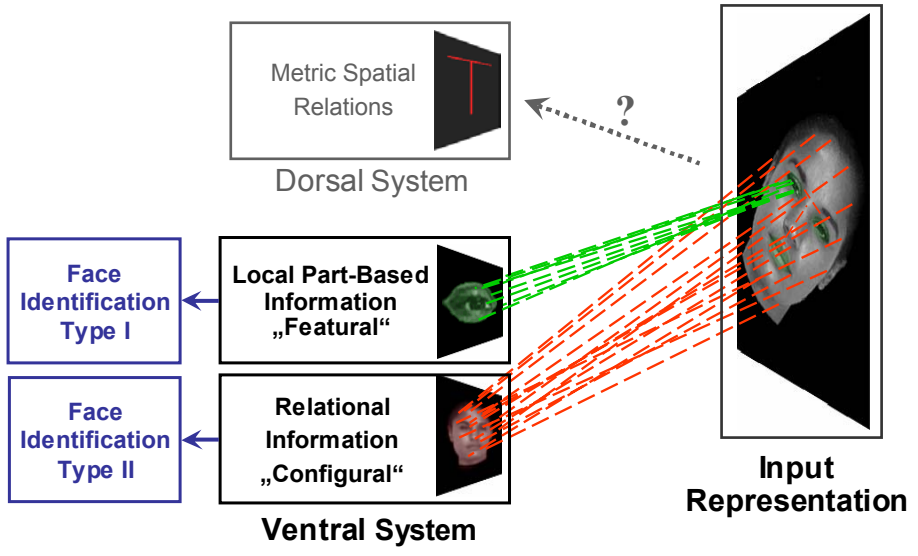
Conclusions Part 2

1. Familiar and unfamiliar face recognition rely on featural and configural information
2. Only quantitative differences.
   = Same relative importance of featural and configural information

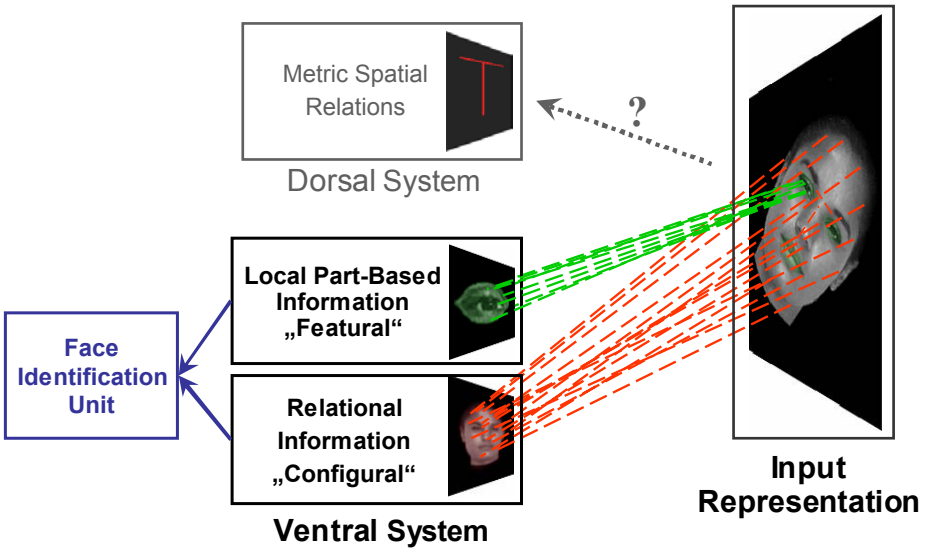
Part 3: Independent or Convergent Processing?



Independent or Convergent Processing?



Independent or Convergent Processing?

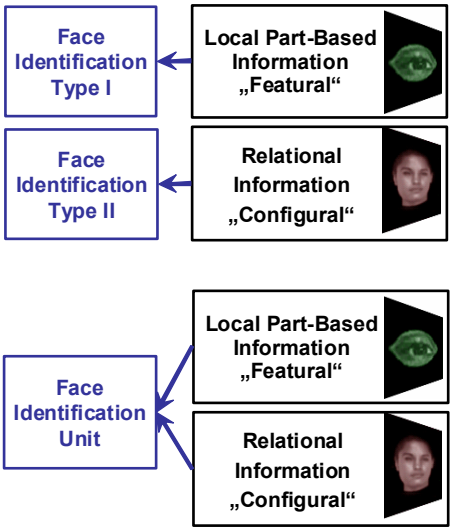


Independent or Convergent Processing?

Method: Repetition Priming

- When identifying objects, people often respond faster the second time an object is shown.
- Repetition priming indicates activation of common representations at the neural level.

Independent or Convergent Processing?



Independent Model

NO Priming

Convergent Model

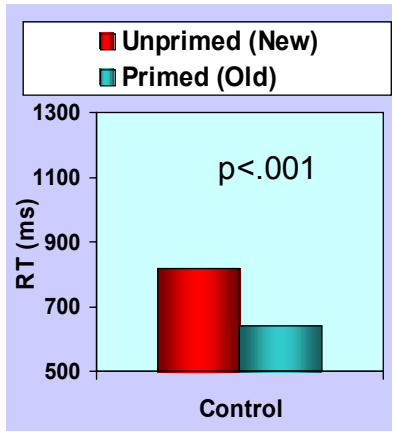
Priming

Control Condition

Group 1
Control
(n=16)

Phase I
10 Celeb-rities

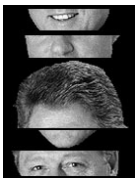
Phase II
10 New
10 Old



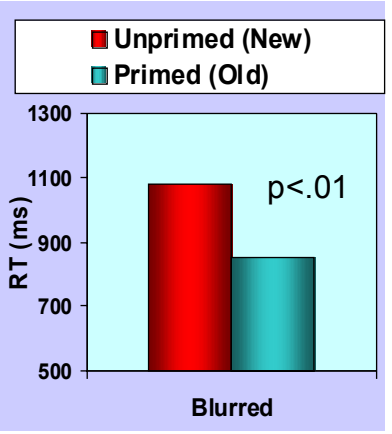
Priming Featural → Configural

Group 2 (n=16)

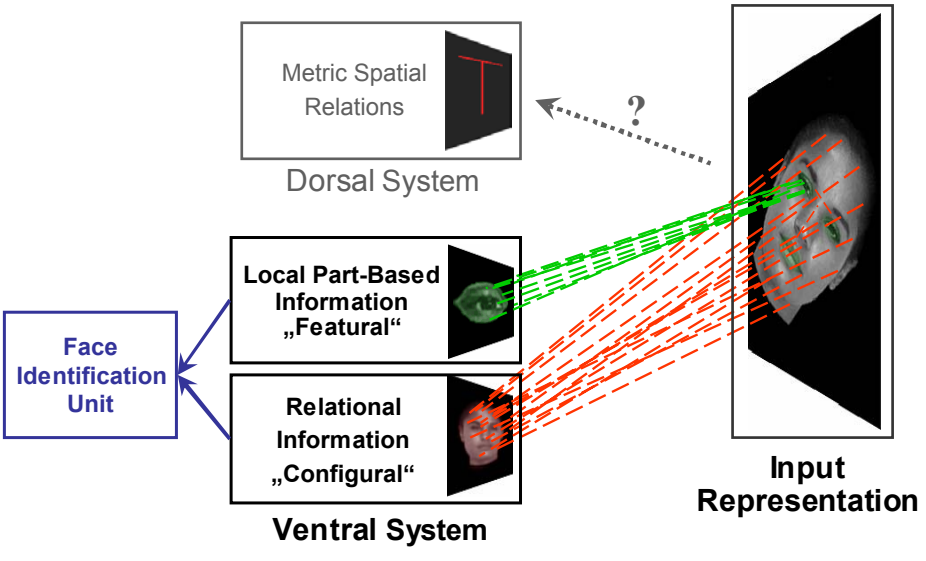
Phase I
10 Celeb-rities



Phase II
10 New
10 Old



Priming Featural → Configural

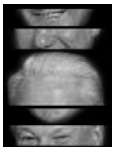


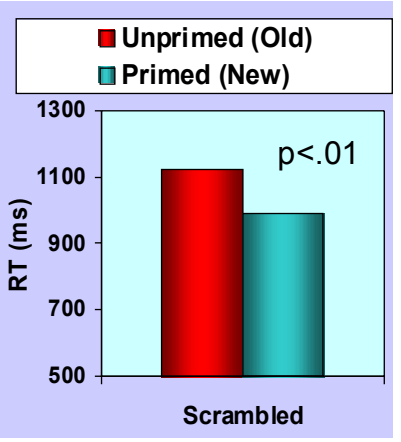
Priming Configural → Featural

Group 3 (n=16)

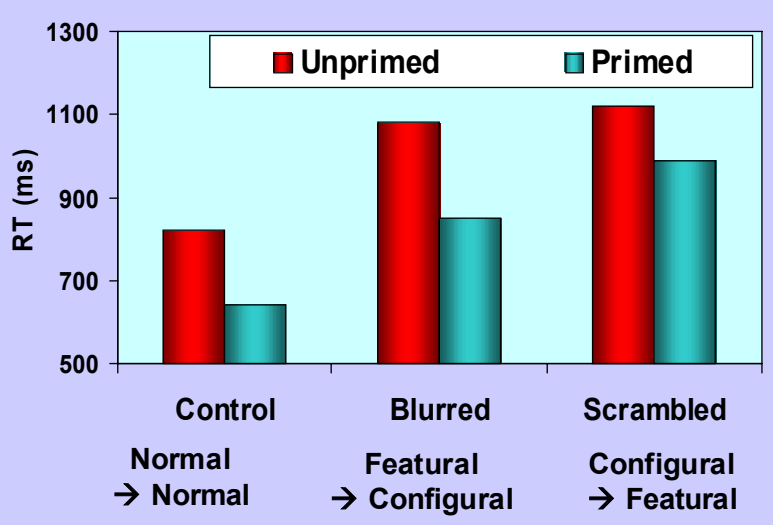
Phase I
10 Celeb-rities

Phase II
10 New
10 Old





Similar Effect of Priming in All Conditions

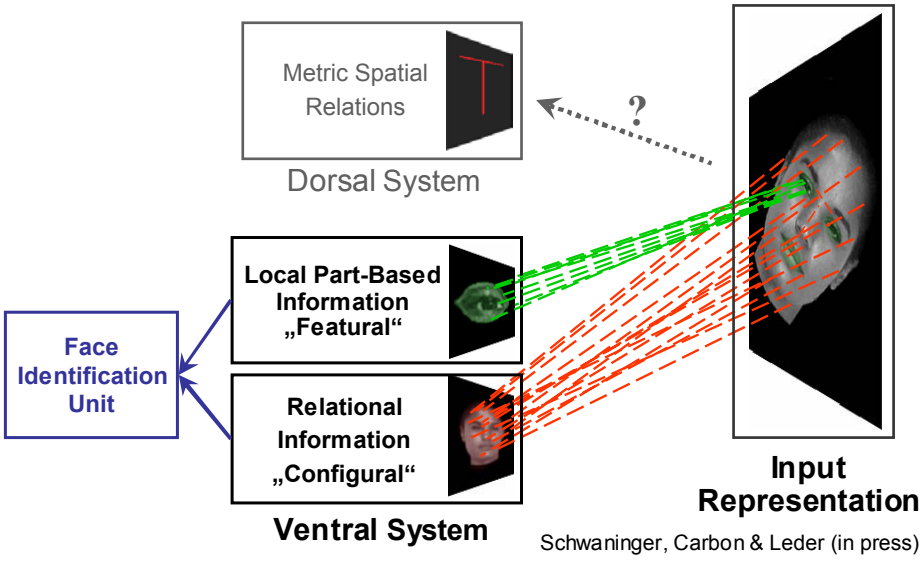


Schwaninger et al. (2002); Collishaw, Schwaninger & Hosie (under review)

Conclusions Part 3

- There are separate explicit representations for featural and configural information.
- These representations can be activated independent of each other.
- The outputs of featural and configural processing converge to the same recognition units.

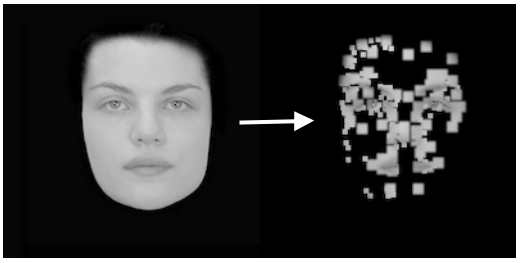
Integrative Modell of Face Recognition



Schwaninger, Carbon & Leder (in press)

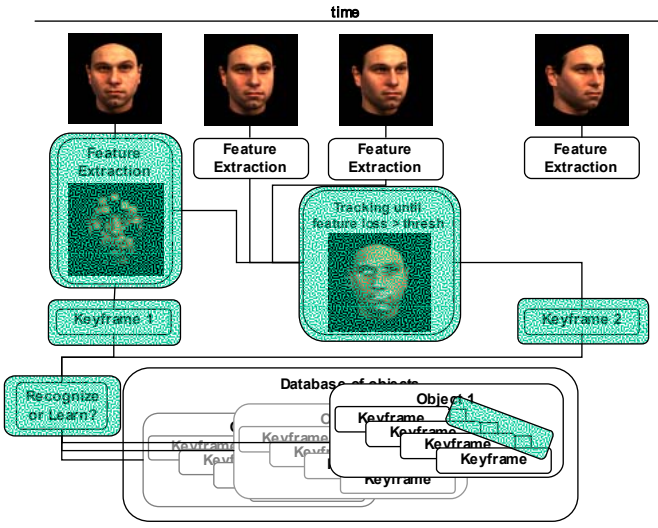
Part 4: Computational Modeling

- **Aim: Implement configural and component processing route**
- Processing of faces is based on **visual features**
  - Detect interest points at coarse and fine scales
  - For each interest point:
    - Store small neighborhood as a **pixel patch**
    - Calculate pixel distances to all other features and store in a **distance histogram**

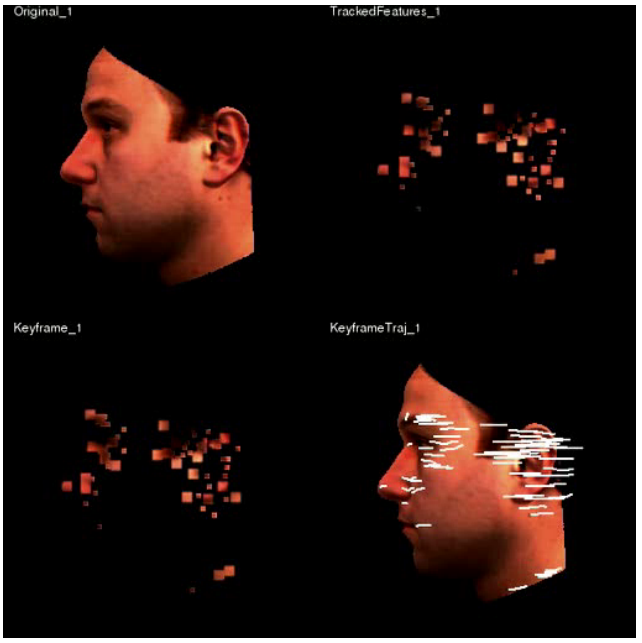


Keyframe framework

Wallraven & Bülthoff 2001



Original images



Tracked features between keyframes

Keyframes

Feature trajectories between keyframes

Matching by correspondences

- How do we match two images?
  - Find corresponding features!
- Matching algorithm constructs a similarity matrix **A**, where each element determines a similarity between feature pairs:

$$A_{ij} = \exp(-\frac{1}{\sigma_{app}^2} app^2(i,j)) \cdot \exp(-\frac{1}{\sigma_{emb}^2} emb^2(i,j))$$

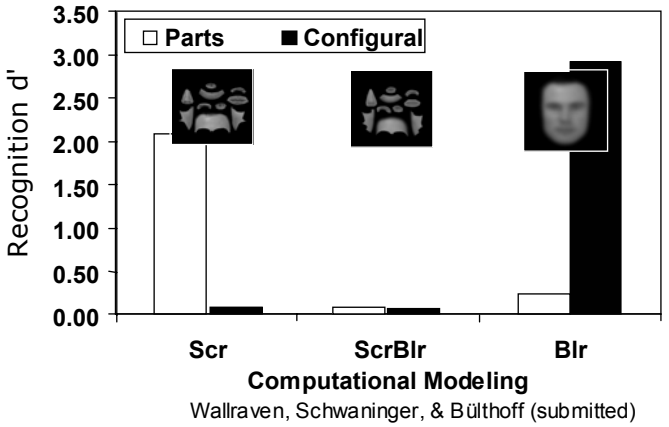
- *app* captures visual similarity (cross-correlation) between two image patches
- *emb* captures the similarity between the distance histograms (χ^2 distance)
- Find largest elements of **A** for corresponding features

Component and configural processing

- Component processing
  - Used at **detailed** scale
  - Calculate *emb* only for distance histograms of the N nearest features
  - Only local embedding of features is used
  - This prefers local clusters of detailed features
- Configural processing
  - Used at **coarse** scale
  - Calculate *emb* for the whole distance histogram
  - Uses global embedding of features
  - This prefers global configurations of coarse features

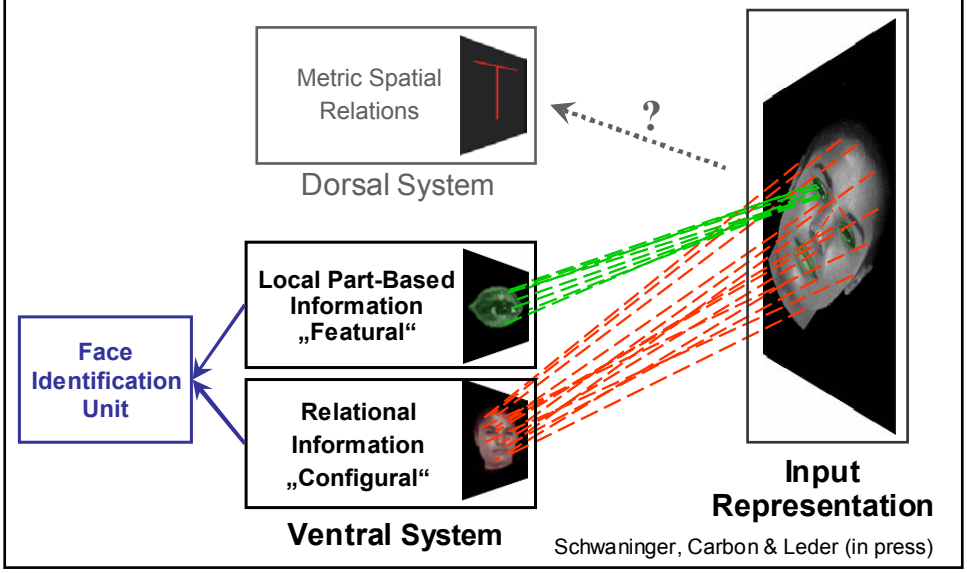
Computational Modeling

- Same stimuli blurred, scrambled and scrambled+blurred condition from Schwaninger et al. (2002)
- Excellent qualitative similarity to human data



Wallraven, Schwaninger, & Bülthoff (submitted)

Integrative Modell of Face Recognition

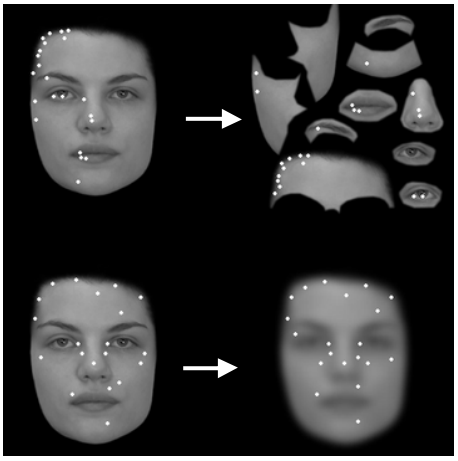


Schwaninger, Carbon & Leder (in press)

Computational Modeling (Example)

Component route in scrambled condition

Configural route in blurred condition

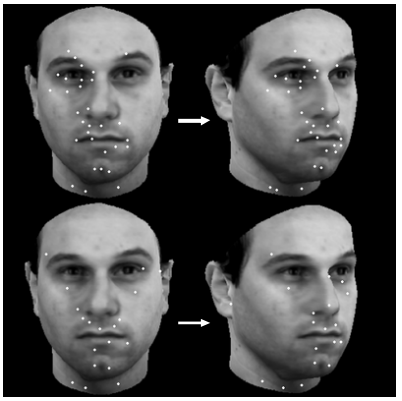


Matching under large view rotations

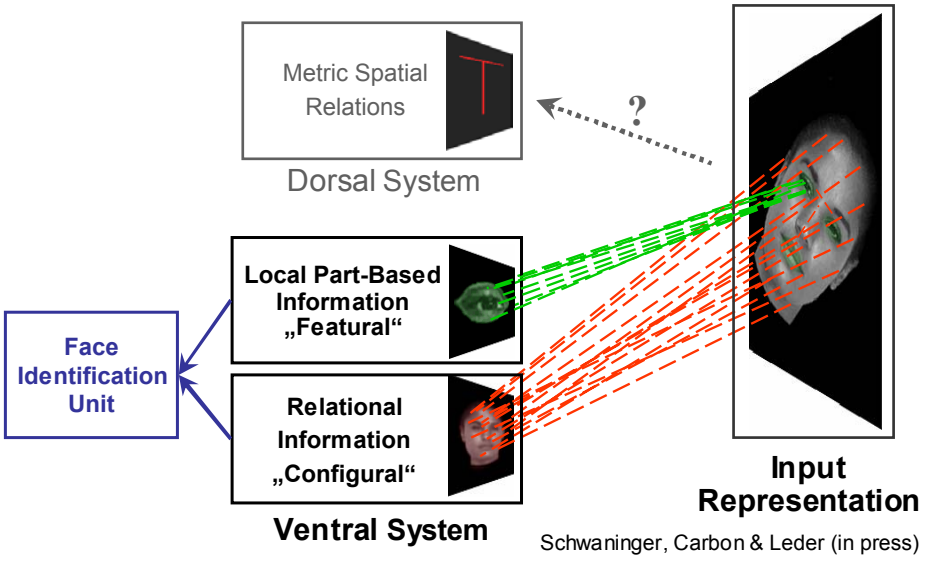
- Configural matching allows matching under large view rotations
- Very good performance when compared to other methods up to 40°!!

Matching using configural route

Matching without configural route

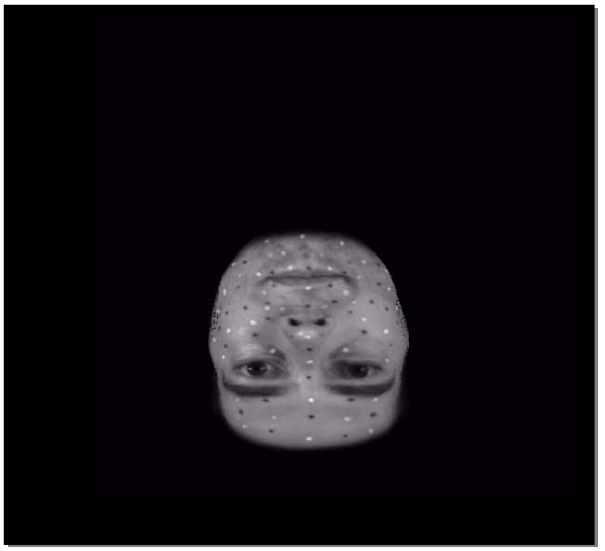


Integrative Modell of Face Recognition

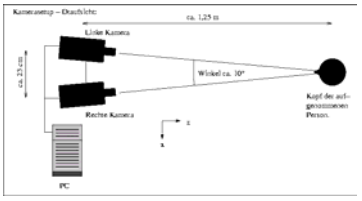


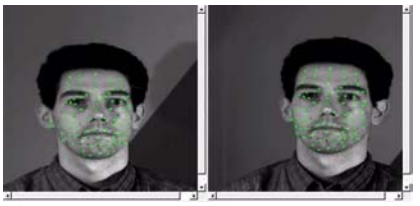
Schwaninger, Carbon & Leder (in press)

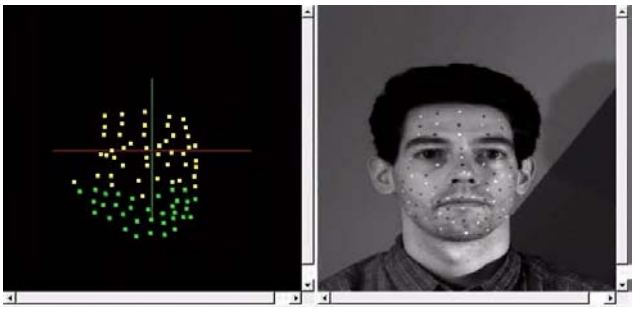
Part 5: Application to Moving Thatcherized Faces



Stimulus Creation: 3D Motion Tracking (Kleiner 2001)

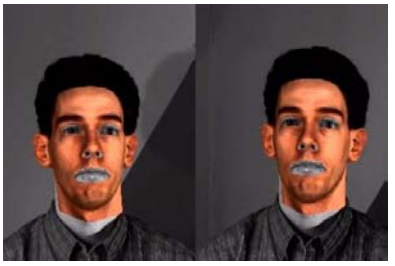






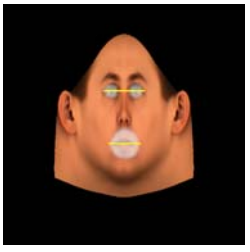
Stimulus Creation: Morphable Shape Model and Texture Extraction

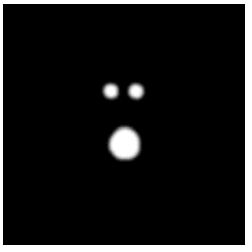
- A morphable 3D shape model (Blanz & Vetter 1999) was fitted to the 3D cloud of tracked points
- Recovered head position and orientation + 3D shape model allowed mapping between single pixels in the video image and corresponding points on the surface (texels) of the persons head.
- Texture map of the filmed face including facial features like eyes, mouth region, nose, forehead etc.

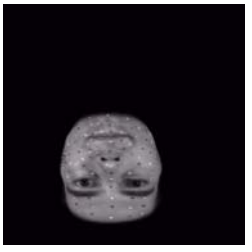


Stimulus Creation: Texture Manipulation & Rendering

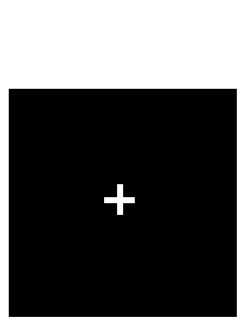
- For „thatcherizing“ the face, the texture content of the eye and mouth regions was automatically mirrored around a horizontal axis in each texture map.
- Smooth transitions between upright and mirrored regions were created by blurring the region boundaries.
- These texture images were reapplied to the persons 3D head model and the head model was rendered (without head motion) in an upright or upside down orientation in front of a black background







Task and Procedure



1 sec Fixation Cross

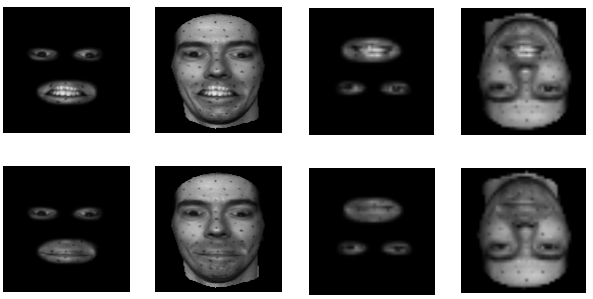
1.5 sec Display

Bizarreness Rating (0-9)

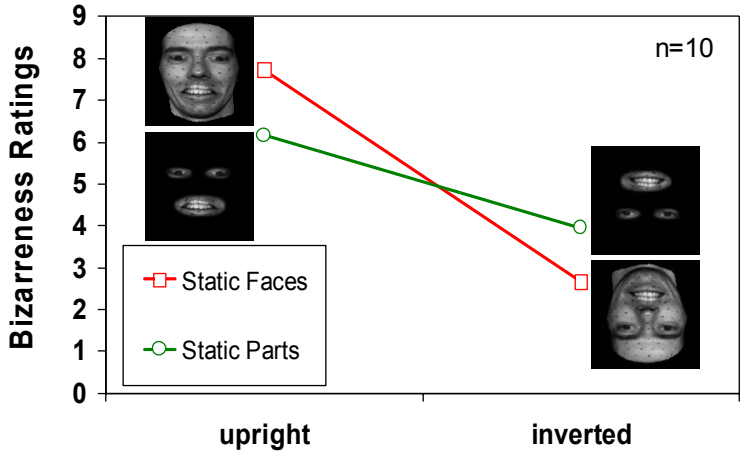
time

Experiment 1: Smiling Faces

- Within-subjects factors:
  - Motion (static vs. moving)
  - Orientation (upright vs. inverted)
  - Info type (parts vs. wholes)
- 2^3* 4(sequences) * 4(rep.) = 128 trials



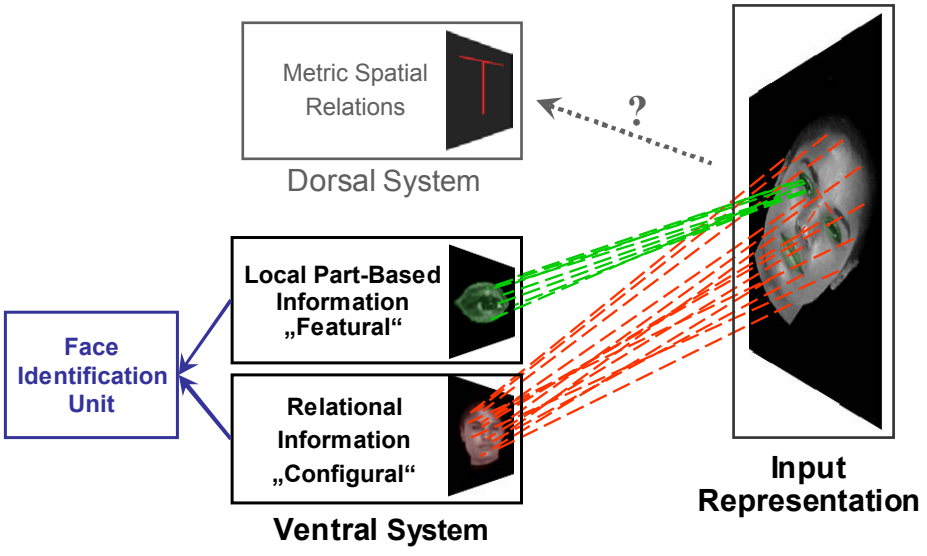
Experiment 1: Smiling Faces



Orientation: $F(1,9) = 41.98$, $MSE = 6.11$, $p < .001$
Orientation * info type: $F(1,9) = 42.69$, $MSE = 2.03$, $p < .01$

Schwaninger, Kleiner & Cunningham (2002)

Integrative Model

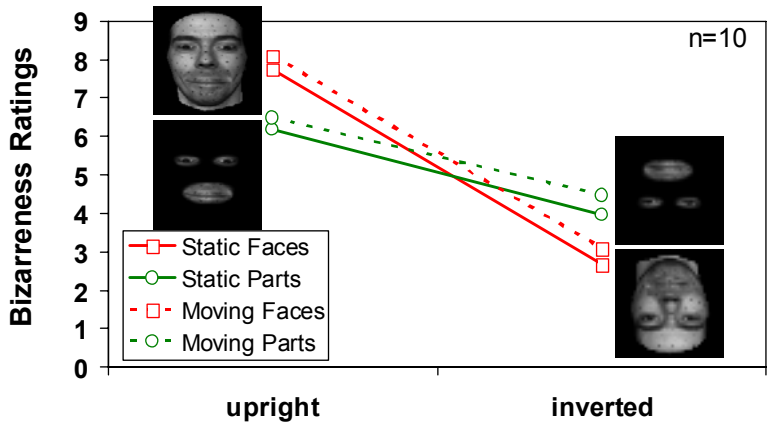


Results Experiment 1: Parts

Interaction between orientation and information type: $p < .01$

- The Thatcher illusion is due to parts and their configuration in upright faces (upright face > upright parts)
- Inverted parts might activate configural representations (basic configuration) (inverted parts > inverted face)

Experiment 1: Smiling Faces



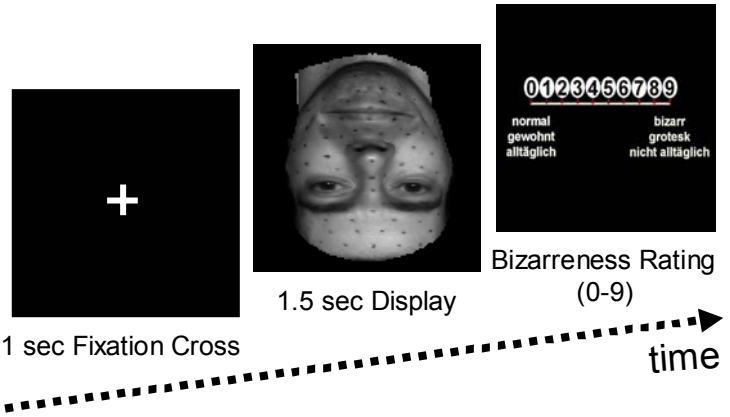
Orientation: $F(1,9) = 41.96$, $MSE = 6.11$, $p < .001$
Orientation * info type: $F(1,9) = 42.69$, $MSE = 2.03$, $p < .01$
Motion: $F(1,9) = 12.59$, $MSE = 0.244$, $p < .01$

Schwaninger, Kleiner & Cunningham (2002)

Results Experiment 1: Motion

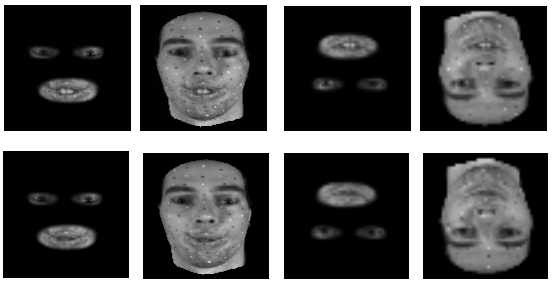
- Motion increases bizarreness
  Main effect of motion: $p < .01$
- This motion effect seems to be local rather than holistic
  - Motion effect independent of facial context:
    Motion * info type: $p = .95$
  - Motion effect independent of orientation:
    Motion * orientation: $p = .603$

Task and Procedure Experiment 2



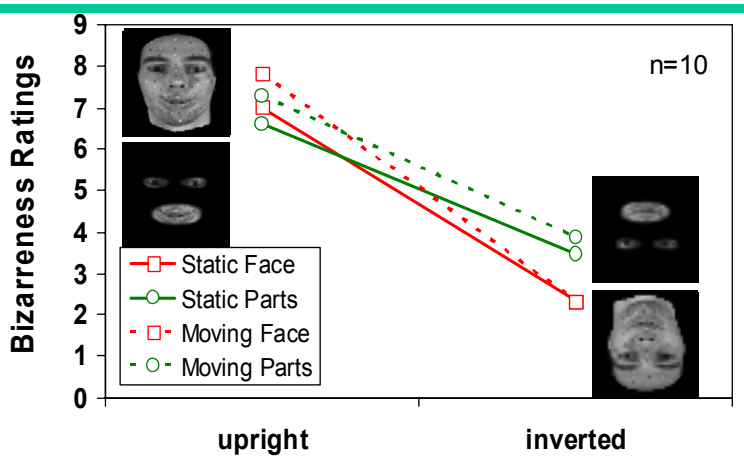
Experiment 2: Talking Faces

- Within-subjects factors:
  - Motion (moving vs. static)
  - Orientation (upright vs. inverted)
  - Info type (parts vs. wholes)
- 2^3* 4(sequences) * 4(rep.) = 128 trials



Experiment 2: Talking Faces

Introduction
1. Rotation
2. Parts / Conf.
3. Model
4. Computation
5. Motion
Summary

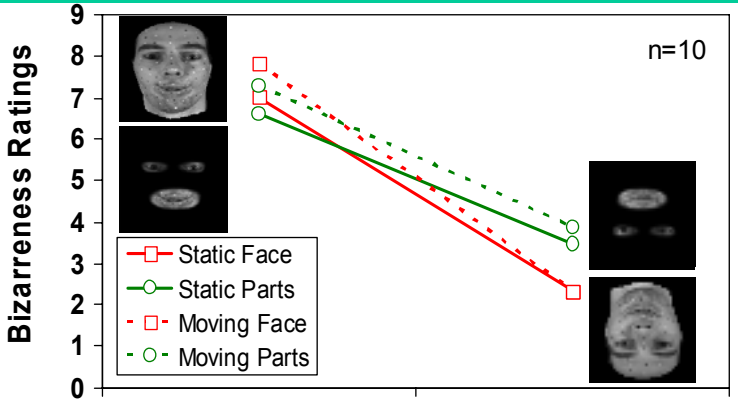


Orientation: $F(1,9) = 185.48$, $MSE = 1.89$, $p < .001$
Orientation * info type: $F(1,9) = 22.10$, $MSE = .74$, $p < .01$
Motion: $F(1,9) = 10.92$, $MSE = 0.40$, $p < .01$
Orientation * Motion: $F(1,9) = 13.89$, $MSE = .11$, $p < .01$

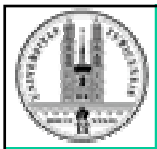
Results Experiment 2: Motion

- Generally, Experiment 2 replicated Experiment 1
- Motion increased bizarreness
  Main effect of motion: $p < .01$
- This motion effect is independent of facial context, i.e. not holistic:
  Motion * info type: $p = .49$
- Type of motion could be relevant
  Exp.1 (smiling): Effect of motion independent on orientation: Motion * orientation: $p = .603$
  Exp. 2 (talking): Motion * orientation: $p < .01$

Experiment 2: Talking Faces

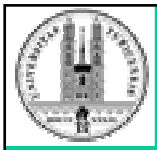


Orientation: $F(1,9) = 185.48$, $MSE = 1.89$, $p < .001$
Orientation * info type: $F(1,9) = 22.10$, $MSE = .74$, $p < .01$
Motion: $F(1,9) = 10.92$, $MSE = 0.40$, $p < .01$
Orientation * Motion: $F(1,9) = 13.89$, $MSE = .11$, $p < .01$



Conclusions Part 5

- The Thatcher illusion is due to parts and their configuration in upright faces (face context increases bizarreness)
- Motion increases bizarreness
- This motion effect is not holistic (independent on face context)
- This motion effect is the same for upright and inverted faces or facial parts



Summary Parts 1-5

- Faces recognition is very orientation-sensitive because faces overtax mental rotation (Part 1)
- Face recognition relies on explicit representations of parts and configural information (Part 2)
- Integrative model for face recognition (Part 3)
- This model is computationally plausible (Part 4)
- Thatcher illusion can be explained by the same model. Motion information is not holistic (Part 5)