Decoding and predicting intentions: A case study for MVPA applications
Decoding intentions

I am going to feed the cat.

Voxel pattern

Pattern vector

Training data

„Feed cat“
„Prepare lecture“
„Book flight“
„Go shopping“
„Post letter“
„Cook dinner“

Test data

Pattern classifier
(e.g. SVM)

Decoded label

„Prepare lecture“

Reverberi, Görgen & Haynes Cer Cor 2011; Bode et al. PLoS One 2011
Task sets as mapping of receptors to effectors
Experimental paradigm “task sets”

Decoding target?  Decoding response?

Cue         Delay     Stimulus          Response
1400ms 2800ms 4200ms 2800ms

Decoding target stimuli?
Decoding mapping?
Decoding response?
Labels for decoding stimuli

Stimulus 1

Stimulus 2

Bode & Haynes Neuroimage 2009
Labels for decoding responses

Response „L“

Response „R“

Bode & Haynes Neuroimage 2009
Labels for decoding task sets

Task set A

Task set B
What to put into your classifier (space)

Whole-brain decoding
Curse of dimensionality (→ PCA)
Biological plausibility?

Region of interest

Searchlight decoding
Specific subspace
Locality assumption
Spatially unbiased
What to put into your classifier (time)

- Single volume activity (typically baseline corrected)
- HRF-fitted single-trial parameter estimate
- Runwise GLM parameter estimates
Spatial normalization can occur before or after classification. This flow chart shows "early" normalization.
Averaging accuracy/information across subjects

One local decoding accuracy map per subject

Average map of local decoding accuracy

FWE-threshold

(T-test, permutation test)

Thresholded map of local decoding accuracy

You can do this in SPM if you are ok with a T-test
Note of caution:
When using a searchlight and a second level threshold these accuracies are not independent estimators of effect size / information (for a related discussion See Vul et al. 2009). For this you would need an additional validation data set. The analysis primarily assesses the existence of information at a specific location.
Decoding of stimulus-response mapping

Bode & Haynes Neuroimage 2009
Decoding Across Time:
Finite Impulse Response Functions

- Canonical HRF
- FIR \rightarrow \text{separate estimate at each TR}
Decoding Across Time:
Finite Impulse Response Functions

Estimate classifier separately for each latency
Decoding across time

Task-Rules
- visCor [-18 -78 0]
  Accuracy = 63%
- IPS [-24 -45 42]
  Accuracy = 60%
- pVLPFC [-48 -12 21]
  Accuracy = 56%
- aVLPFC [-51 33 3]
  Accuracy = 58%

Target Stimuli
- visCor [-15 -9 3]
  Accuracy = 72%
- aVLPFC [-51 33 -12]
  Accuracy = 56%

Motor Responses
- SMA [-18 6 66]
  Accuracy = 65%
- motCor [-39 15 57]
  Accuracy = 75%

Bode & Haynes Neuroimage 2009
Hierarchy and compositionality: Cross-classification

Rule AB

Compositional Coding

Rule A

Independent Coding

Rule B

CUE, 500 ms “if there is a face press left”

DELAY, 3500-9500 ms

TARGET, 750 ms

DELAY, 2000-4000 ms

CUE, 500 ms “if there is a house press left” “if there is a face press right”

Reverberi, Görgen & Haynes (Cer Cortex 2012)
Hierarchy and compositionality: Cross-classification

DLPFC [45 24 48] accuracy = 58%

VLPFC [33 51 -12] accuracy = 59%

Reverberi, Görgen & Haynes (Cer Cortex 2012)
Predicting decisions

... potential predictors ...

Soon, Brass, Heinze & Haynes Nature Neurosci 2008
Decoding at various TRs relative to choice

Soon, Brass, Heinze & Haynes Nature Neurosci 2008
Decoding choice after decision

Soon, Brass, Heinze & Haynes Nature Neurosci 2008
Decoding choice before decision

Lateral frontopolar cortex

Medial frontopolar cortex

Precuneus / posterior cingulate cortex
Searchlight classification weights in frontopolar cortex
Early decision – late response?

Lateral FPC

Decoding accuracy vs. Time [s]

Left motor cortex

Decoding accuracy vs. Time [s]
No prediction of next trial

Increases with distance from previous trial
Sequential dependencies in detail

Allefeld et al. Arxiv 2013; see also Lages et al. 2013
Which model? Two “labelling” approaches

Soon et al. Frontiers (2014)
Pitfalls: Reaction times

One possible "fix": Regress out reaction times before classification

Todd et al. Neuroimage 2013
Pitfalls: Statistical testing

Especially for chance accuracies other than $p=0.5$ better to use permutation tests.

Görgen & Haynes (in preparation)
Group members:
Carsten Allefeld, Stefan Bode, Carsten Bogler, Thomas Christophel, Radolaw Cichy, Kerstin Hackmack, Kai Görgen, Martin Hebart, Jakob Heinzle, Fatma Imamoglu, Adrian Imfeld, Thorsten Kahnt, Christian Kalberlah, Ida Mommenejad, Fernando Ramirez, Carlo Reverberi, Chun Siong Soon, Anita Tusche, Martin Weygandt, David Wisniewski

Collaborators:
Benjamin Blankertz, Matthias Schultze-Kraft, Tobias Donner, Aaron Schurger