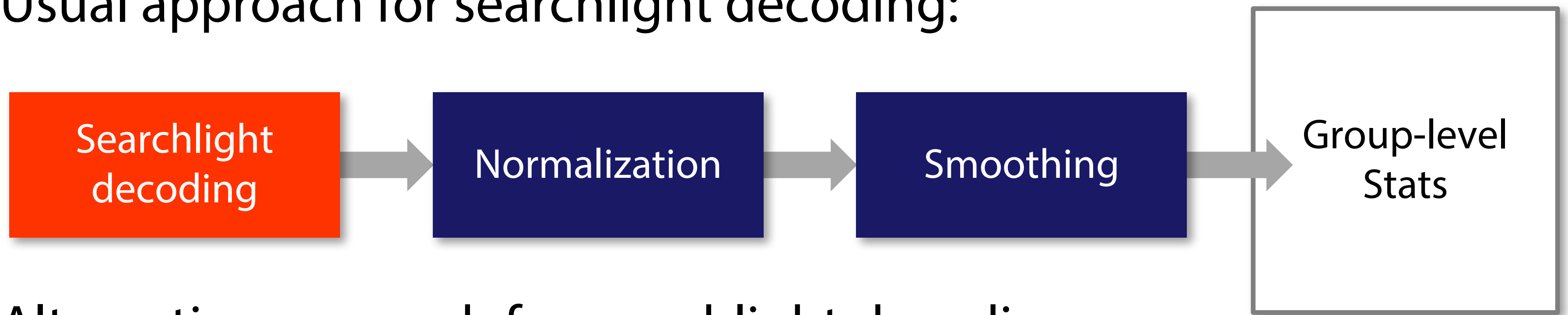
Session IV: Design of MVPA Studies

Martin Hebart
Laboratory of Brain and Cognition
NIMH

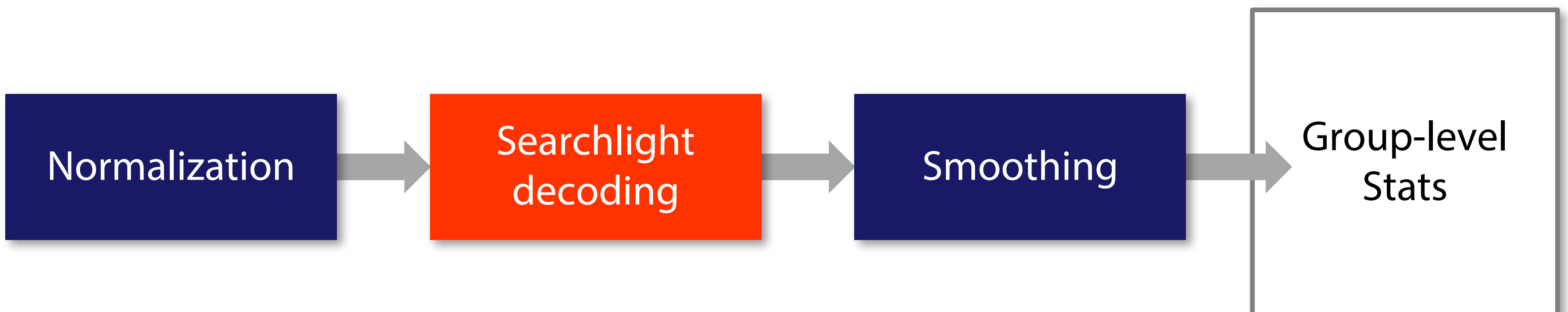
BUT BEFORE WE GET TO EXPERIMENTAL DESIGN…

General fMRI Preprocessing

Usual approach for searchlight decoding:



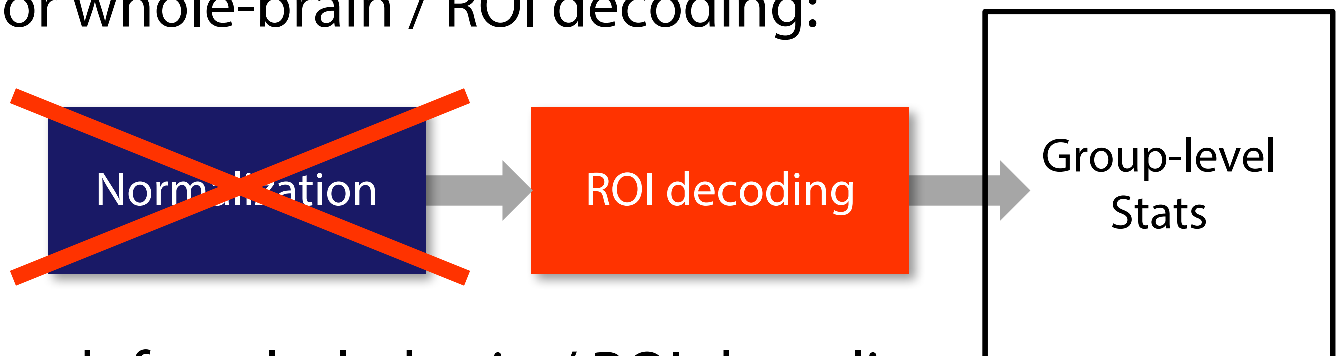
Alternative approach for searchlight decoding:



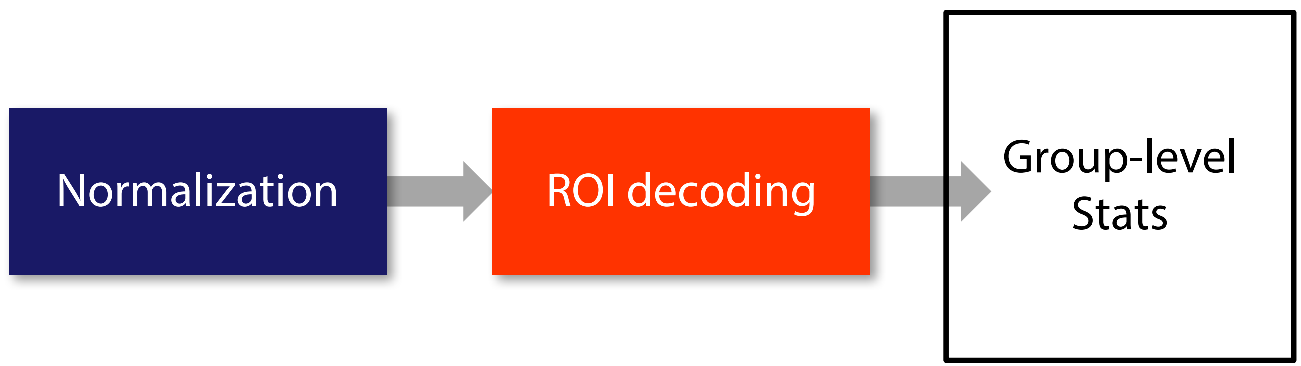
Second approach required for between-subject analysis

General fMRI Preprocessing

Usual approach for whole-brain / ROI decoding:



Alternative approach for whole-brain / ROI decoding:



➡ Second approach required for between-subject analysis

EXPERIMENTAL DESIGN

Goals of this Presentation

MVPA Design

- What are requirements for the experimental design of MVPA studies?

Minimize Noise, Maximize Signal

- How can we maximize the information we extract by our experimental design?

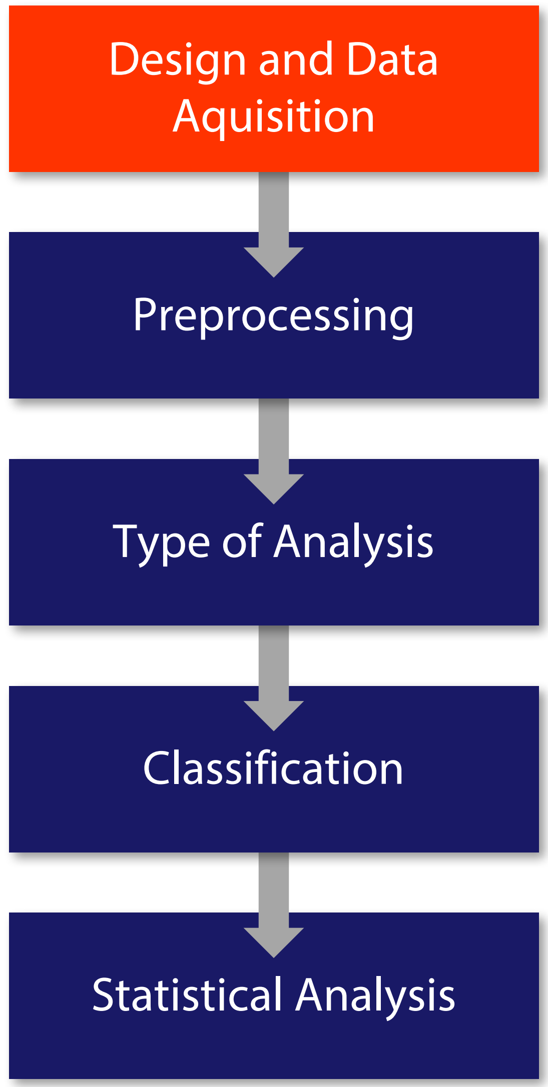
Confounds in MVPA studies

- What are possible confounds that we have to consider?
- How can we avoid these confounds?
- If we cannot avoid them: How can we detect and eliminate them?

Important Note

- How you later want to analyze your data can have important consequences for your design
- This presentation: strong focus on fMRI classification within subject (e.g. different conditions)
- <u>But</u>: Principles apply also to between subject classification or other modalities (e.g. EEG)

MVPA Workflow



MVPA DESIGN: GENERAL

Design choices for Multivariate Decoding

Most design choices identical to univariate GLM

- duration of experiment (longer = better)
- scanner settings (TR, TE, flip angle, descending acquisition, …)
- high spatial resolution: unclear if benefit specific to multivariate

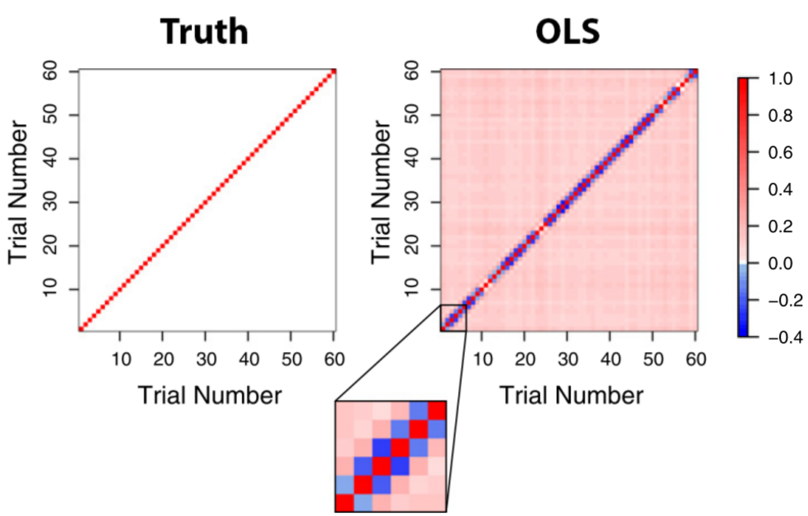
What you need to consider

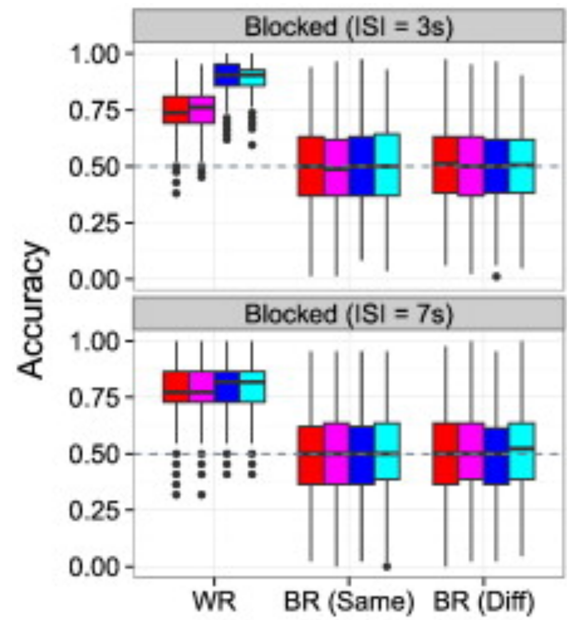
- classification requires independence (or dependence balanced between classes) of training and test data
- often data dependencies exist that invalidate this assumption
- **crucial for cross-validation: <u>all</u> conditions needed in <u>all</u> folds**

Leave-One-Run Out Cross-Validation

Typical Analysis: Leave-one-run-out cross-validation

Reason: Non-independence within run can bias results





Mumford et al (2014) – Neuroimage

How Many Runs for Leave-One-Run Out?

	Many short runs	Few long runs
Data variability	more variable	more stable
Amount of training data	more training data	less training data

- Usually 6 to 12 runs, 4 runs is usually minimum
- Number often determined by condition balancing within run
- When minimizing between-run differences: more runs better

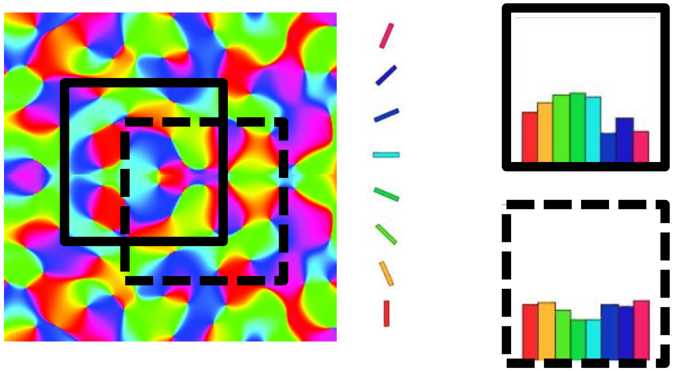
Coutanche & Thompson-Schill (2012) – Neuroimage

HOW TO MINIMIZE NOISE AND MAXIMIZE SIGNAL

Minimize Noise

- Assuming that source of information lies in fine-scale patterns, noise perturbations can destroy this information
- Most important sources of noise that affect fine-scale patterns:
  - head motion
  - physiological noise
  - scanner-related effects

Example: Effect of displacement on sampling of orientation columns



→ motion correction only interpolates

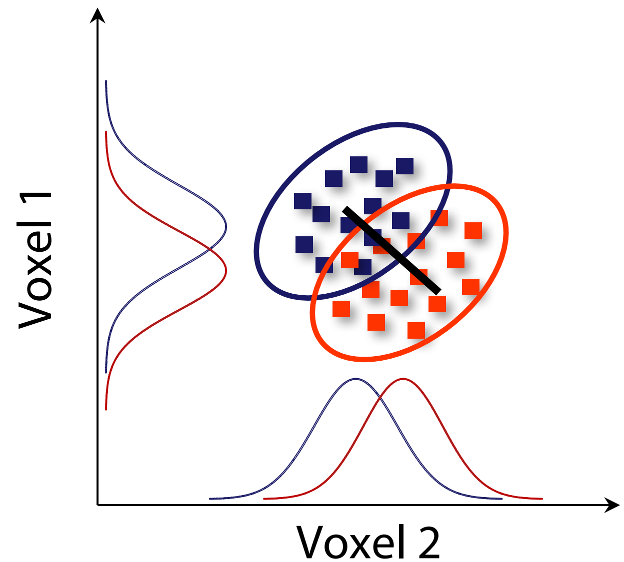
Swisher et al (2010) – J Neurosci; Freeman & Heeger (2011) – J Neurosci

Maximize Signal: Design Efficiency

Design Efficiency

Goal: Maximize the pattern distinctness, i.e. increase between-class scatter or decrease within-class scatter

→ Improving design efficiency of GLM for univariate contrast will maximize the distinctness also for MVPA



Additional advice:

For single-trial event-related analysis and short TR (< 2.5s), time-locking onsets to TR can reduce between trial variability

http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency

Interim Summary

- Reducing noise in the acquisition may have a stronger benefit for multivariate analyses than univariate analyses, but benefit is still unclear
- Optimizing the design efficiency can improve the pattern distinctness
- Common software for doing this at the end of the presentation

THE PERVASIVENESS OF CONFOUNDS IN MVPA STUDIES

Confounds

Two classes of confounds:

1. Confounds in the experimental design
2. Confounds in the results

→ Best practice is to avoid confounds before they happen

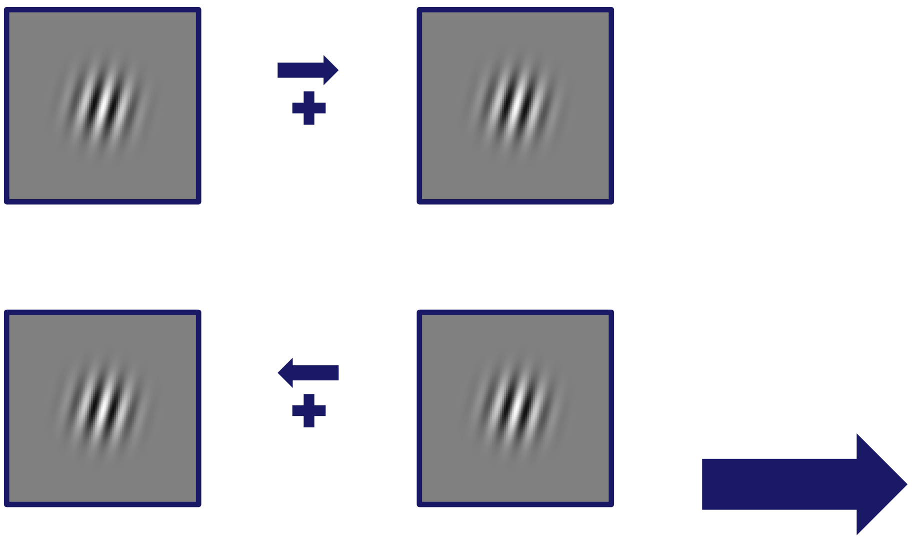
→ We can avoid confounds in experimental design

→ There are some confounds we cannot avoid, but can only control

Confounds in the Experimental Design

Typically disregarded issues can become a confound

Example: Classification of top-down visual attention

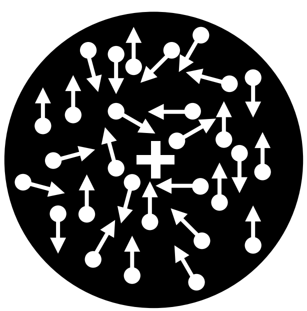


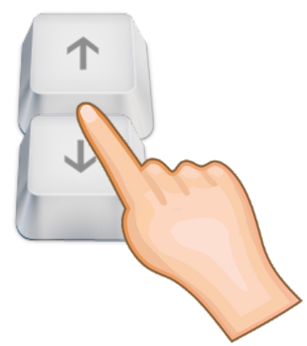
Confound: visual cue

Confounds in the Experimental Design

Typically disregarded issues can become a confound

Example: Classification of choice?





Confound: motor response

Confounds in the Experimental Design

Solution 1: Exclude brain regions that confound responds to

- For motor confound: Exclude motor-related brain regions
- For visual confound: Exclude visually responsive regions

<u>But</u>: Maybe unexpected brain region responds to confound? Maybe those regions respond to true effect?

<u>And</u>: Sometimes approach not possible

→ not recommended

Confounds in the Experimental Design

Solution 2: Separate confound in time or through jitter

- For motor confound: Wait 20s with motor response
- For visual confound: Show cue jittered and model separately

<u>But</u>: Pattern autocorrelation can last very long

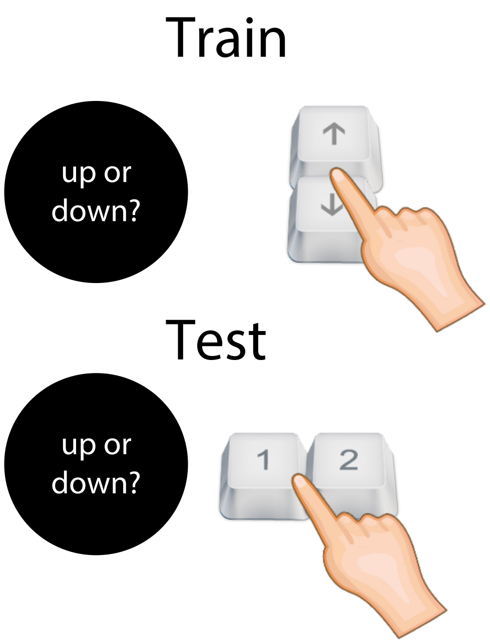
<u>And</u>: Jitter only reduces confound, never eliminates it!

➡ not recommended

Confounds in the Experimental Design

Solution 3: Cross-classification

- For motor confound: Switch response modality, e.g. after each run
- Train classifier on data with one confound, test on data with other
- If above chance, then classifier generalizes across confound
- For visual confound: Use different cue



<u>But</u>: Less data available for training, i.e. possibly reduced sensitivity

<u>And</u>: Possible task-switching costs

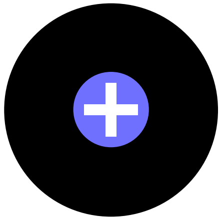
→ recommended if no better possibility

Confounds in the Experimental Design

Solution 4: Cue Trick

- For motor confound: Use response-mapping rule
- Controls for confound by balancing
- For visual confound: Not always possible

Rule 1

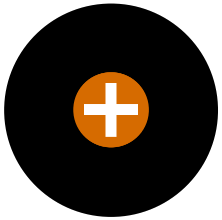


if up,
press left
if down,
press right

<u>But</u>: Requires training of subject

→ recommended when possible

Rule 2

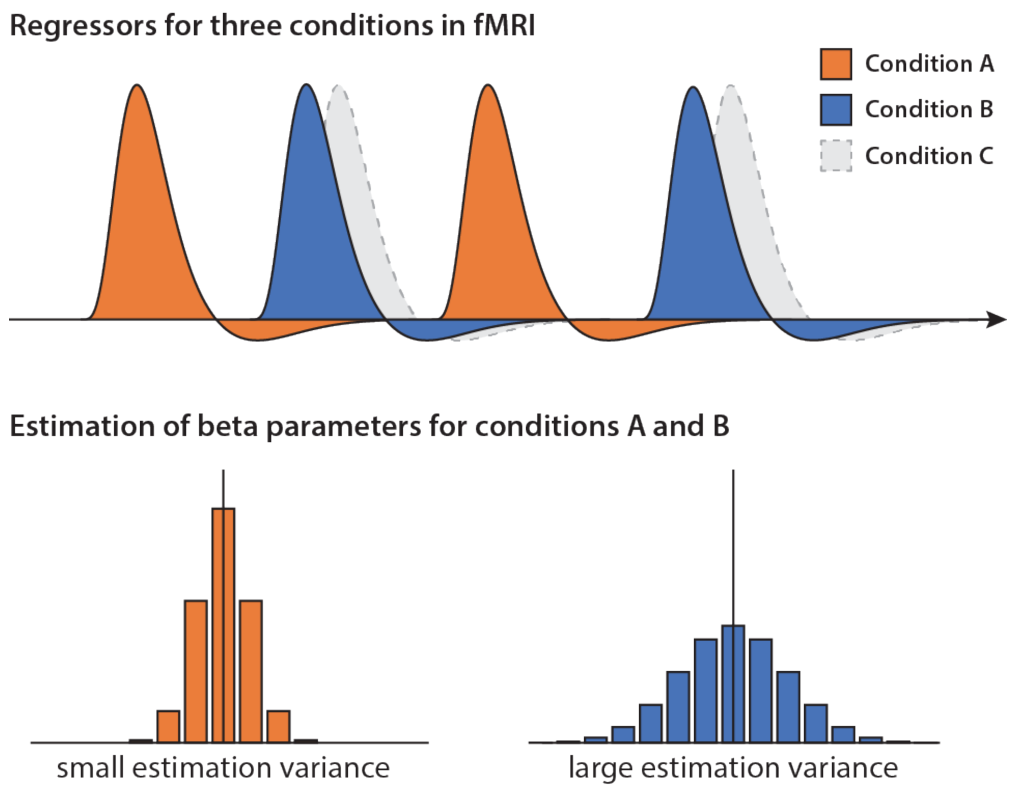


if up,
press right
if down,
press left

Hebart et al (2014) – Cerebral Cortex

Confounds in the Experimental Design / Results

Different estimability of betas for class A vs. B problematic



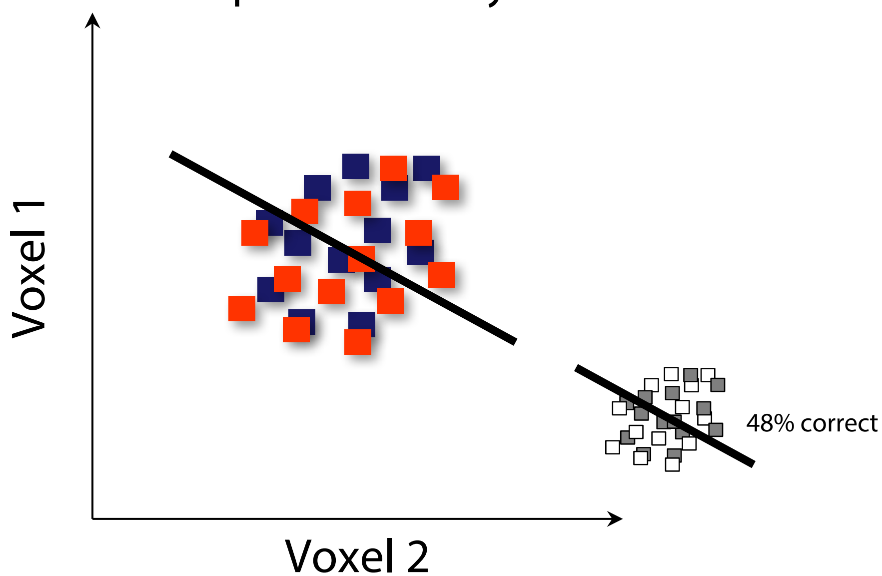
Hebart & Baker (2017) – Neuroimage; Görgen et al (2017) – bioRxiV

Confounds in the Experimental Design / Results

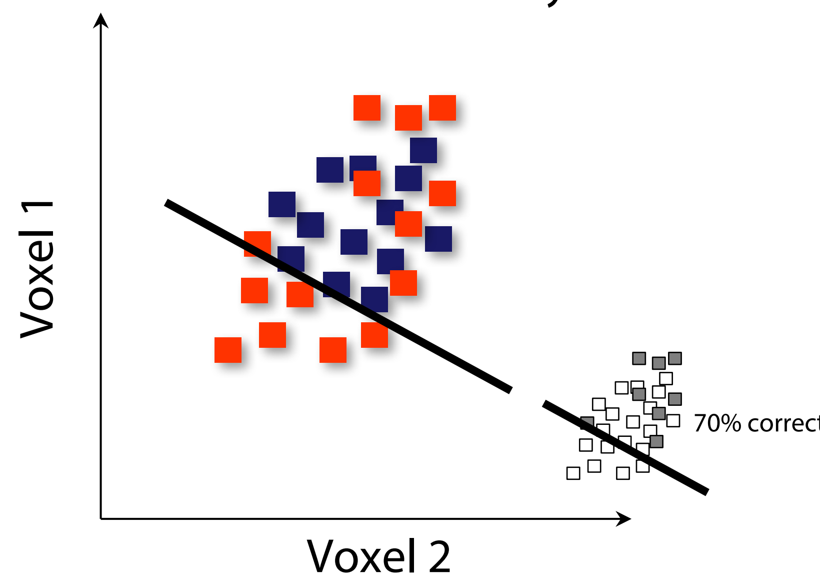
Different estimability of betas for class A vs. B problematic

Result without any difference in mean pattern

equal variability



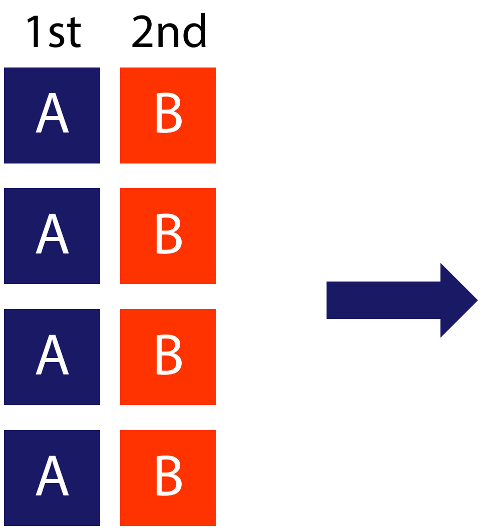
different variability



Hebart & Baker (2017) – Neuroimage; Görgen et al (2017) – bioRxiV

Detection of Confounds

Example: Order confound



Impossible to distinguish if effect:

1. true difference in A vs. B

or

2. based on fatigue (A first, B later)

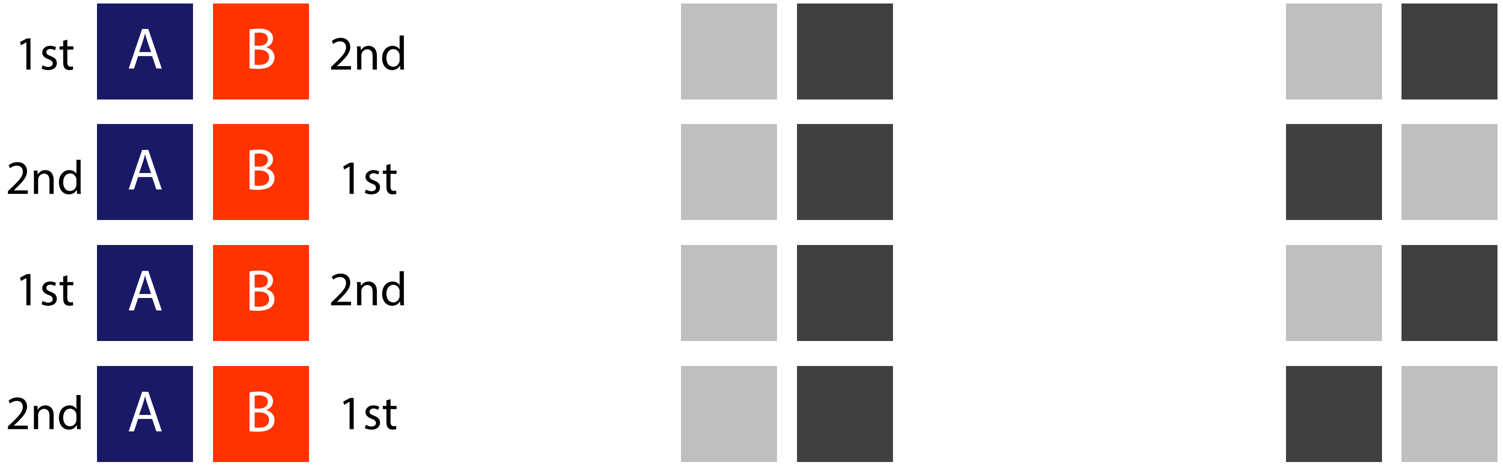
Görgen, Hebart, Allefeld, & Haynes (2017) – arXiv

Detection of Confounds

Possible solution to order confound: Counterbalancing?

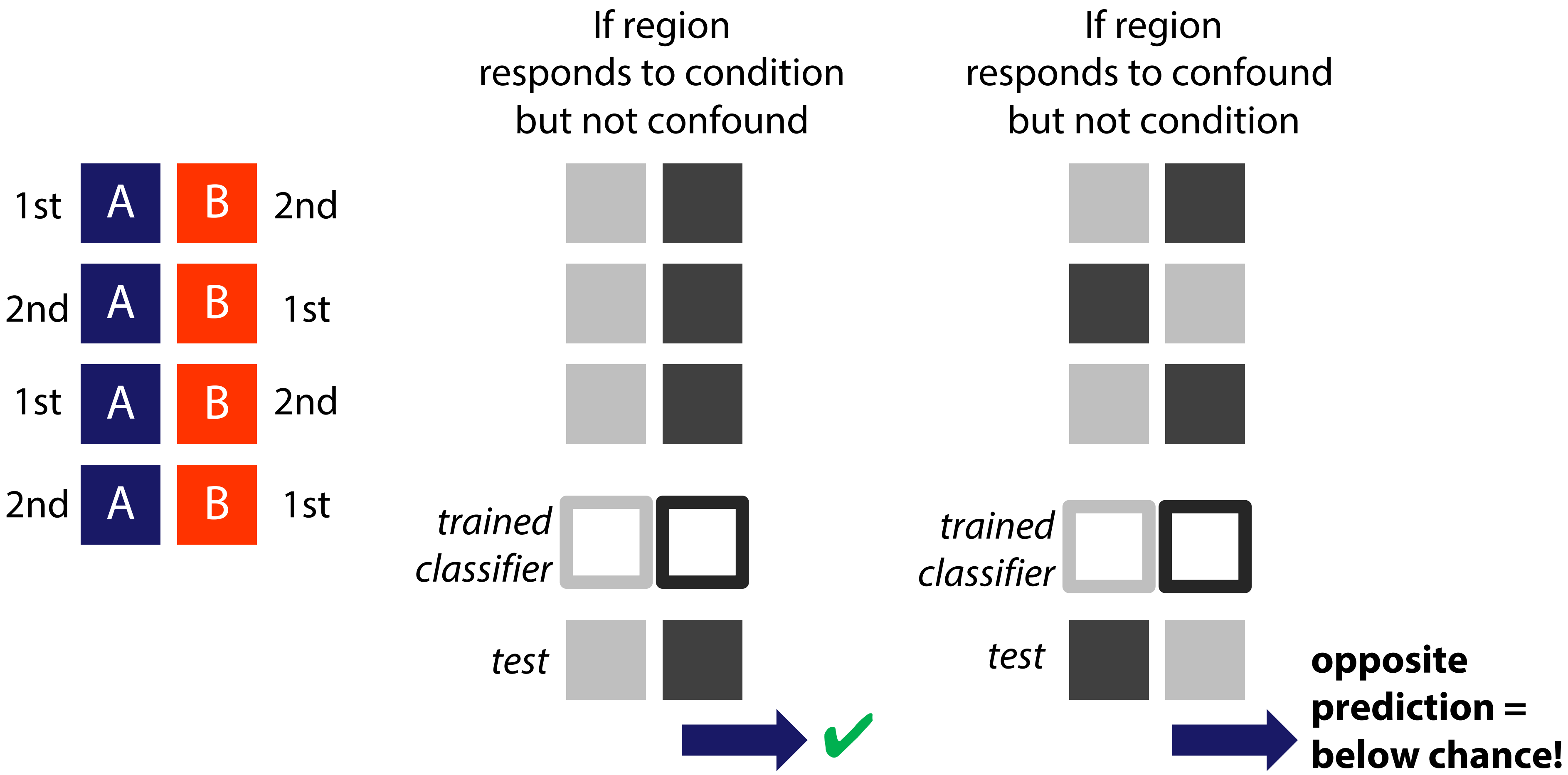
If region responds to condition but not confound

If region responds to confound but not condition



Detection of Confounds

Possible solution to order confound: Counterbalancing?



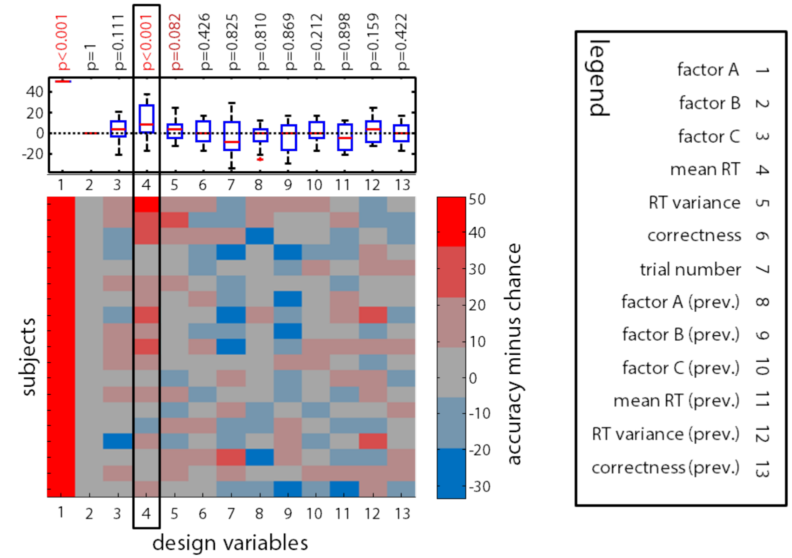
Detection of Confounds in Design

Same analysis approach (Görgen et al., arXiv)

- We do decoding analysis using cross-validation
- Cross-validation is a different statistical method than classical statistics
- To measure the influence of a confound, we need to apply the same statistical method to it

→ Univariate decoding on confounding variable, i.e. treat confound as data

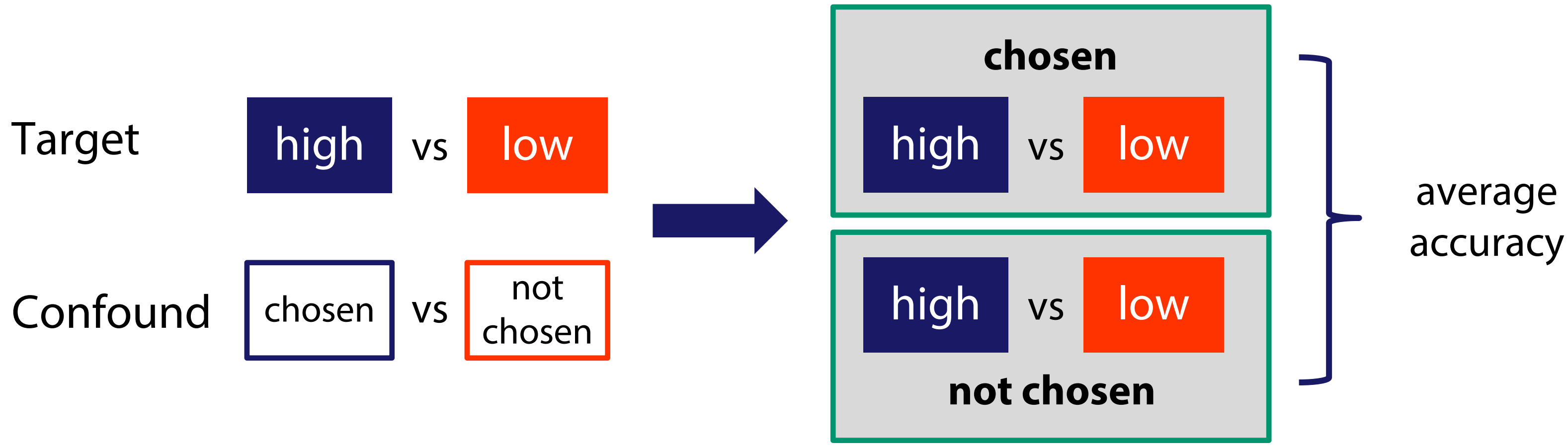
→ Can become part of efficiency optimization if confound in design is assumed



Elimination of Confounds

Approach 1: Balance confounds

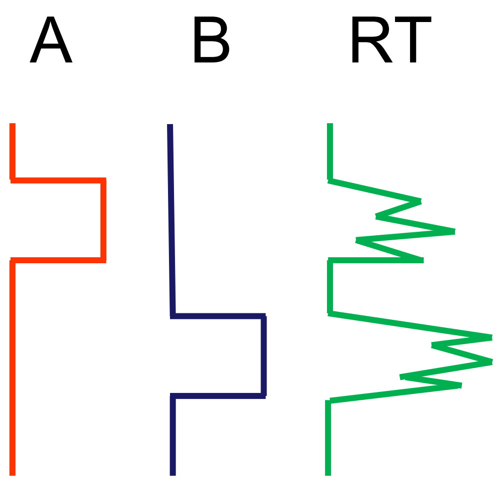
- Do sub-classification on data
- Example: Representation of value in choice task



<u>but</u>: less sensitive than one classification (ameliorated by ensemble approach with repeated subsampling)
<u>and</u>: difficult to apply for continuous confounding variables

Elimination of Confounds

Approach 2: Add nuisance regressor to remove confound



- can introduce bias when control regressors correlated differently between classes
- can lead to false sense of certainty that confound has been controlled

Todd et al (2013) – Neuroimage; Woolgar et al (2014) – Neuroimage

Summary

- Experimental design requires all conditions to be roughly equally present in all folds (e.g. runs)
- Confounds are difficult to deal with
- The cue trick is useful to decorrelate choices and button presses
- It is not easy to follow the own intuition to avoid confounds
- Below-chance accuracies (and false above-chance accuracies) can be explained by uncontrolled or badly-controlled confounds
- The same analysis approach provides a method for detecting confounds before they occur

Study Question

(1) If you are planning an MVPA study or have MVPA data, think about possible confounds in your experimental design. How would you deal with them? Discuss them in your small group.

(2) Alternatively, design a very simple experiment where participants make choices to two different stimuli.

(a) How can you avoid confounding choices with button presses?

(b) Can you think of other confounds? How can you take these into account?

Maximize Signal

Software for Design Efficiency

- Doug Greve: OptSeq

  http://surfer.nmr.mgh.harvard.edu/optseq/

- Wager & Nichols: Genetic algorithm

  https://github.com/canlab/CanlabCore

- No optimization algorithm, but easy and more flexible method to set up design matrix:

  http://martin-hebart.de/webpages/code/matlab.html

  Suggestion: Brute-force repetition works well in general