Voodoo correlations
and double dipping: a cautionary tale

Chris Baker

Section on Learning and Plasticity
Laboratory of Brain and Cognition
Circular analysis in systems neuroscience: the dangers of double dipping

Nikolaus Kriegeskorte, W Kyle Simmons, Patrick S F Bellgowan & Chris I Baker

Voodoo Correlations in fMRI Studies of Emotion, Personality, and Social Cognition

Edward Vul,1 Christine Harris,2 Piotr Winkielman,2 & Harold Pashler2

1Massachusetts Institute of Technology and 2University of California, San Diego

HISTORICAL NEWS & VIEWS: REPRODUCIBILITY

Double-dipping revisited

Robust conclusions require rigorous statistics. In 2009 a seminal paper described the dangers and prevalence of double-dipping in neuroscience. Ten years on, I consider progress toward statistical rigor in neuroimaging.

Katherine S. Button
Take Home Message

Data selection should be independent from any other data testing
Faces

Regions specialized for face processing?

Occipital face area

Fusiform face area
High-resolution imaging reveals highly selective nonface clusters in the fusiform face area

Kalanit Grill-Spector, Rory Sayres & David Ress
Example 1
Univariate activation

Non-Independent  Independent

Non-Independent

Independent

**Non-Independent**

**Independent**

**Nonbrain ROI**
Double dipping can produce an effect where there is none

Can also distort a true underlying effect
Simulation:
Regional activation analysis

- 3D voxel volume (30 x 30 x 20 voxels)
- Block-design experiment
- 4 conditions (A, B, C, D)
- Spatiotemporal noise:
  - Gaussian, slightly spatially smoothed
contrast hypothesis

true effects

fMRI signal

A B C D condition

central slice
p<0.05 (corrected)

contrast hypothesis

independent-data ROI

same-data ROI

p<0.01
independent-data ROI

same-data ROI

blending continuum

truth

hypothesis
Simulation: Regional activation analysis

• Double dipping can produce distortions of the data even when there is a true underlying effect
Example 2

Multivariate Pattern Analysis (MVPA)
Example 2: MVPA

Simmons, Martin et al. 2006

TASK
(property judgment)

Animate/Inanimate? Pleasant/Unpleasant?

STIMULUS
(object category)
Example 2: MVPA

Simmons, Martin et al. 2006
Split-half analysis (Haxby et al., 2001)

• Define ROI
  – Select voxels for which any pairwise condition contrast is significant

• Correlation analysis
  – Split data in half (training and test)
  – Compare activity patterns for same versus different conditions
Decoding accuracy

fMRI data

Task

Stimulus

0.5

Chance level

0

1
But training and test were independent.

All data used to select voxels

Half data used to select voxels

Decoding accuracy

fMRI data

Random generator data

Task

Stimulus

Decoding accuracy

0.5

Chance level 0.5

0

0

0

0
But training and test were independent....

All data used to select voxels

Half data used to select voxels
But training and test were independent....

All data used to select voxels

Half data used to select voxels

Decoding accuracy

fMRI data

Random generator data

Task

Stimulus

Chance level

0.5

0

1

0.5

0.5

0

1

0
Example 2:
Multivariate Pattern Analysis (MVPA)

• Test data must not be used in
  – Training the classifier
  – Defining the ROI
Example 3

Brain-behavior correlations (voodoo)
The Neural Basis of Loss Aversion in Decision-Making Under Risk

Sabrina M. Tom, Craig R. Fox, Christopher Trepel, Russell A. Poldrack

---

**Behavioral loss aversion [ln(λ)]** vs. **Neural loss aversion [(-β_{loss}) - β_{gain}]**

$r = 0.85, P < 0.001$

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Number of voxels</th>
<th>Anatomical location</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>284</td>
<td>L inferior/middle frontal</td>
</tr>
<tr>
<td>0.88</td>
<td>175</td>
<td>R inferior/middle frontal</td>
</tr>
<tr>
<td>0.87</td>
<td>104</td>
<td>L inferior frontal (opercular)/anterior insula</td>
</tr>
<tr>
<td>0.86</td>
<td>122</td>
<td>R inferior frontal (opercular)</td>
</tr>
<tr>
<td>0.85</td>
<td>332</td>
<td>B ventral striatum</td>
</tr>
<tr>
<td>0.83</td>
<td>358</td>
<td>R inferior parietal</td>
</tr>
<tr>
<td>0.81</td>
<td>110</td>
<td>B pre-supplementary motor area</td>
</tr>
<tr>
<td>0.46</td>
<td>963</td>
<td>L lateral occipital/cerebellum</td>
</tr>
</tbody>
</table>

Non-Independent
Double dipping can inflate brain-behavior correlations
Data analysis results
data → analysis → assumptions → results
Circular inference

- data
- assumptions
- analysis
- results
How do assumptions in fMRI produce circularity?

- Selection
Selective Analysis

• Powerful and essential tool
  – Increasingly large datasets
• Common throughout neuroscience
How do assumptions in fMRI produce circularity?

• Selection
  – When the results statistics are not independent of the criteria used for selection

• Why?
  – Selection is based on true effects + noise
Types of selection

• Binary selection
• Continuous selection (weighting)
• Sorting
Examples of selection

• fMRI
  – ROI or voxels of interest
Double dipping is not just a problem for fMRI
Examples of selection

- **fMRI**
  - ROI or voxels of interest

- **Single-unit recording**
  - “typical” neuron
  - pre-specified selectivity e.g. visually responsive

- **EEG/MEG**
  - Sensors or waveforms of interest

- **Behavior**
  - Sorting participants by performance e.g. fast learners versus slow learners

- **Gene microarrays**
  - Subsets of genes
How common is double dipping?

• All fMRI articles published in 2008 in
  – Nature
  – Science
  – Nature Neuroscience
  – Neuron
  – Journal of Neuroscience
How common is double dipping?

134 papers

- 42% contained non-independent selective analysis
- 14% insufficient methodological detail

These papers may not be wrong in their main claims, but it’s very difficult to know how much of a bias has been introduced.
How common is double dipping?

Today?
Data selection should be independent from any other data testing