Advances in Causal Inference Workshop, UAI 2015
July 16, 2015

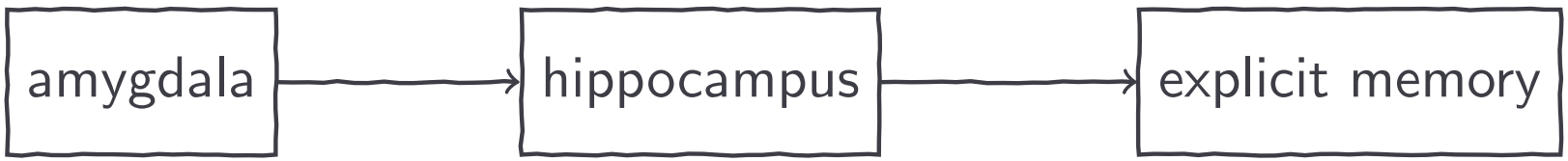
Causal interpretation rules for encoding and decoding models in neuroimaging

Sebastian Weichwald, Timm Meyer, Ozan Özdenizci[§],
Bernhard Schölkopf, Tonio Ball[‡], Moritz Grosse-Wentrup
MPI for Intelligent Systems, [§]Sabanci University, [‡]University of Freiburg

---

sweichwald.de/neuroimage2015  brain-computer-interfaces.net

Motivation



Hippocampal activity in this study was correlated with amygdala activity, supporting the view that the amygdala **enhances** *explicit memory by* **modulating** *activity in the hippocampus.*

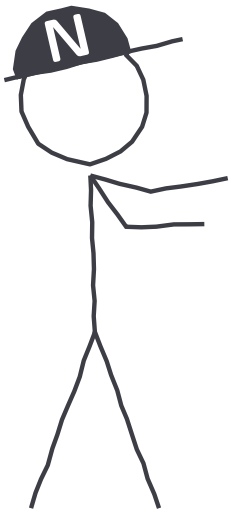
(S. Hamann, *Trends in Cognitive Sciences*, 2001)

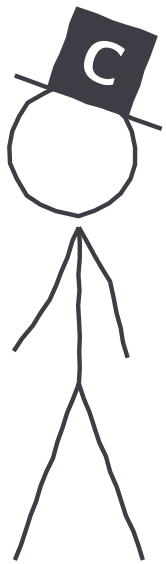
We tested [...] whether pre-stimulus alpha oscillations measured with electroencephalography (EEG) **influence** *the encoding of items into working memory.*

(Myers et al., *Journal of Neuroscience*, 2014)

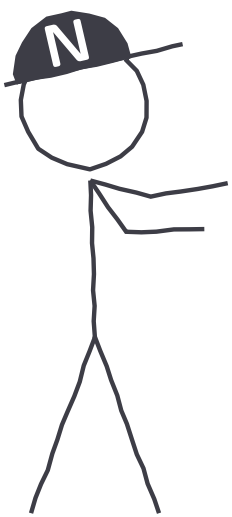
1. Motivation
2. Approach
3. Encoding and decoding models in neuroimaging
4. Causal interpretation of encoding and decoding models
5. Empirical example
6. Wrap-up

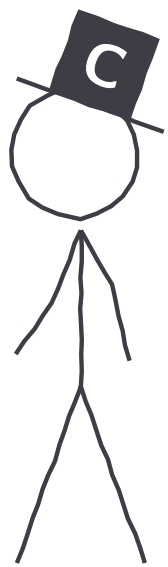
Approach


N


C

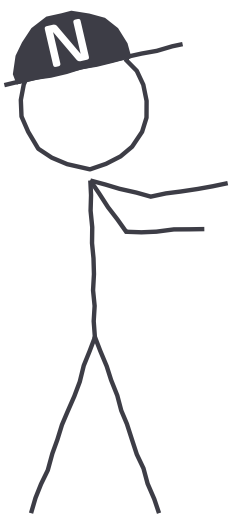
I'm interested in how neural activity gives rise to cognition.

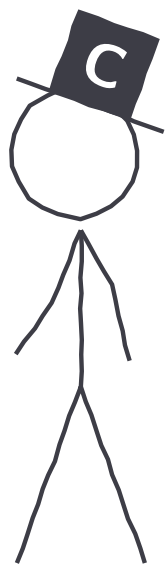




I'm interested in how neural activity gives rise to cognition.

That sounds intriguing! So, what do you do?

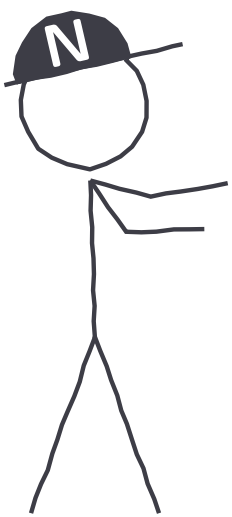


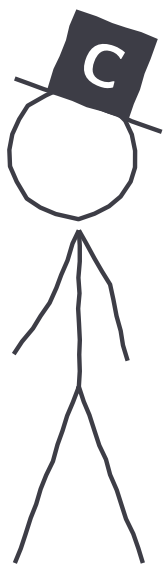


I'm interested in how neural activity gives rise to cognition.

That sounds intriguing! So, what do you do?

I present stimuli to subjects or observe their behaviour while recording their brain activity. ...[explains common analysis methods]...



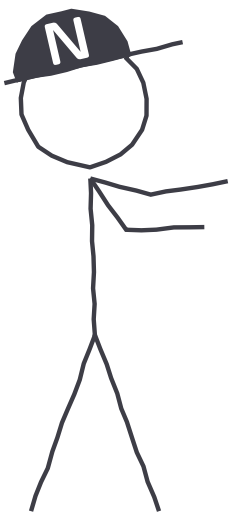


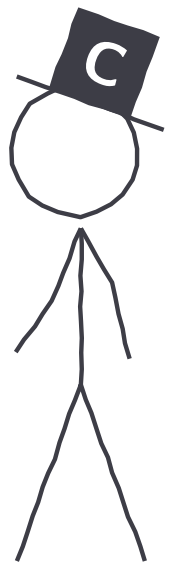
I'm interested in how neural activity gives rise to cognition.

That sounds intriguing! So, what do you do?

I present stimuli to subjects or observe their behaviour while recording their brain activity. ...[explains common analysis methods]...

Ha, interesting! ...[rephrases what has been said in causal inference slang]...





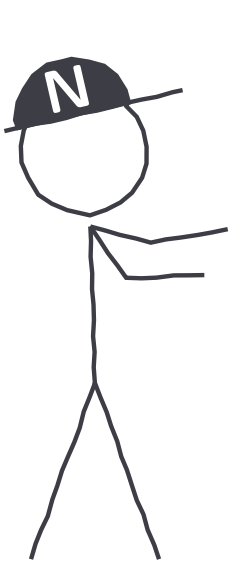
I'm interested in how neural activity gives rise to cognition.

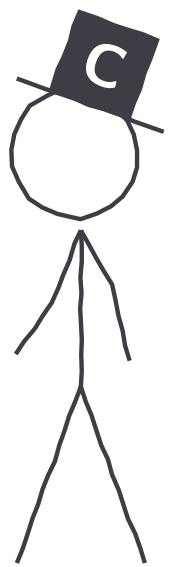
That sounds intriguing! So, what do you do?

I present stimuli to subjects or observe their behaviour while recording their brain activity. ...[explains common analysis methods]...

Ha, interesting! ...[rephrases what has been said in causal inference slang]...

Yeah, a solid basis for our interpretations! It also clarifies problems that we recently discussed in the community.





Encoding and decoding models in neuroimaging

Trial 3

L

Trial 4

L

Trial 5

R

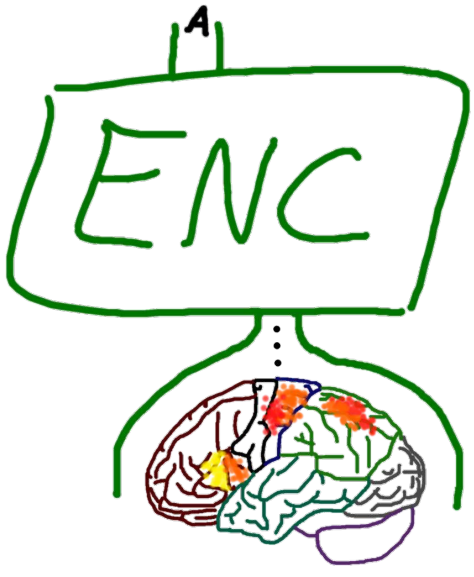
Trial 6

R

Trial

R

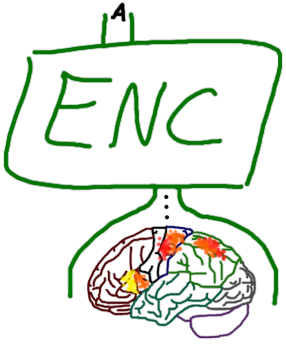
encoding



e.g. mean difference between conditions

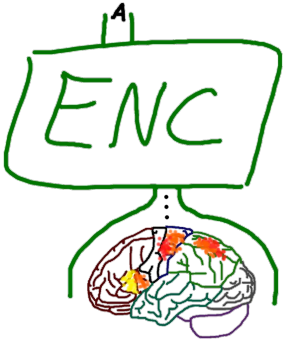
decoding

e.g. classifier for experimental conditions



"Feature shows significant variation across experimental conditions?"

"Feature helpful for predicting the experimental condition?"



"Feature shows significant variation across experimental conditions?"

"Feature helpful for predicting the experimental condition?"

?

relevant feature ⇜⇝ cognitive process

Causal interpretation of encoding and decoding models

Let's set out the causal component of already performed analyses..

Let's set out the causal component of already performed analyses..

stimulus- vs response-based

feature relevance $\not\rightsquigarrow$ marginal/conditional dependence

$\rightsquigarrow$ 16 causal interpretation rules

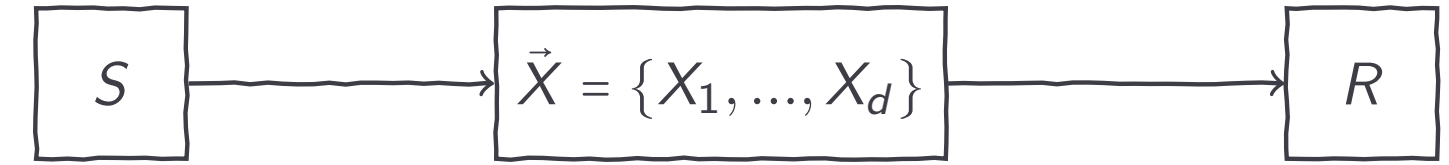
Let's set out the causal component of already performed analyses..

stimulus- vs response-based

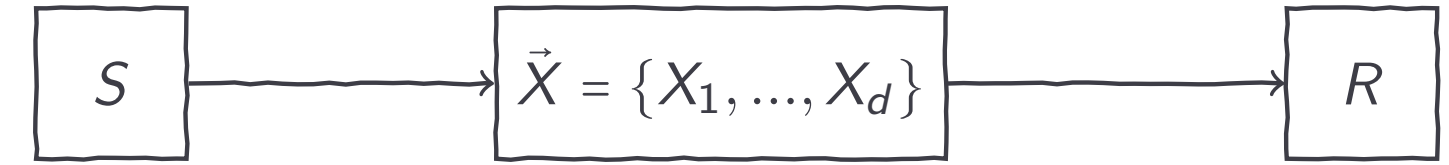
feature relevance $\nsim$ marginal/conditional dependence

$\rightsquigarrow$ 16 causal interpretation rules

simple

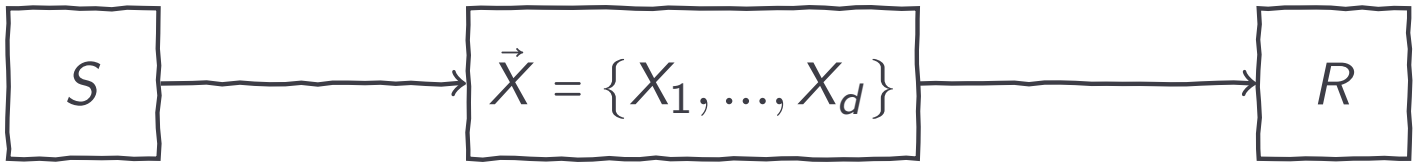


stimulus brain state features response



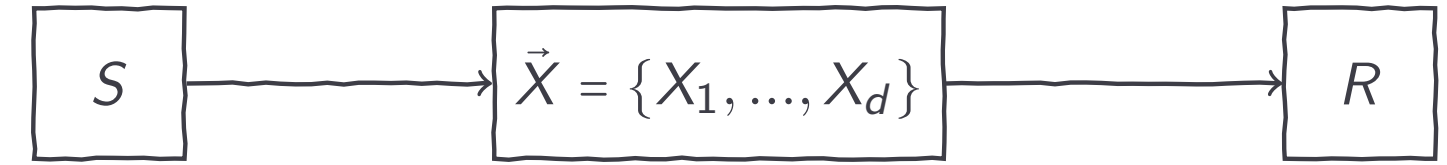
stimulus | brain state features | response

stimulus-based		response-based
$p(\vec{X}\|S)$	encoding	$p(\vec{X}\|R)$
$p(S\|\vec{X})$	decoding	$p(R\|\vec{X})$



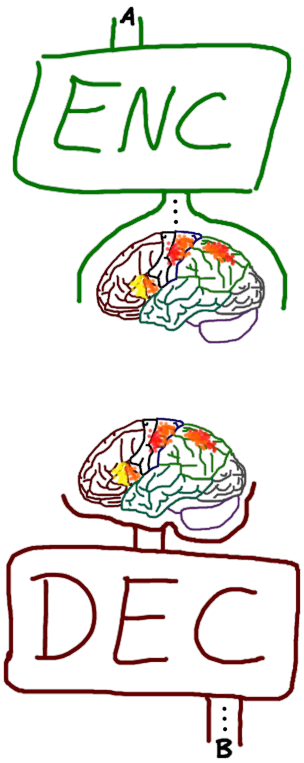
stimulus | brain state features | response

stimulus-based			response-based	
$p(\vec{X}\|S)$	causal	encoding		$p(\vec{X}\|R)$
$p(S\|\vec{X})$		decoding	causal	$p(R\|\vec{X})$

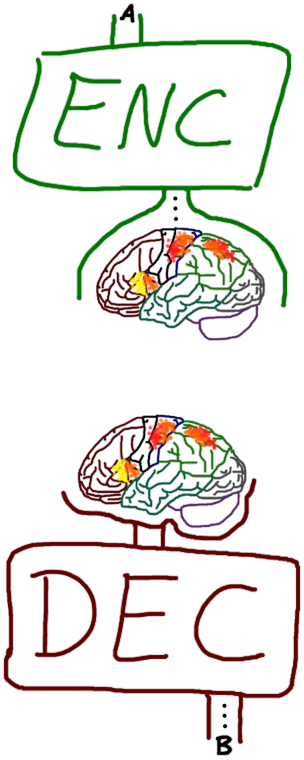


stimulus brain state features response

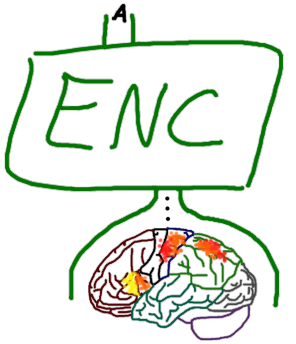
	stimulus-based		response-based	
$p(\vec{X}\|S)$	causal	encoding	*anti*-causal	$p(\vec{X}\|R)$
$p(S\|\vec{X})$	*anti*-causal	decoding	causal	$p(R\|\vec{X})$


A
ENC
DEC
B

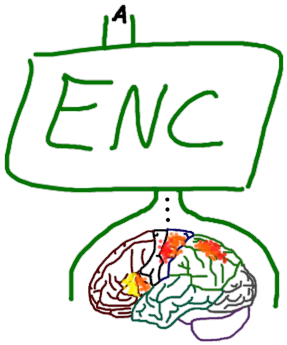

C



$$p(X_i|C = c_1) \stackrel{?}{\neq} p(X_i|C = c_2)$$

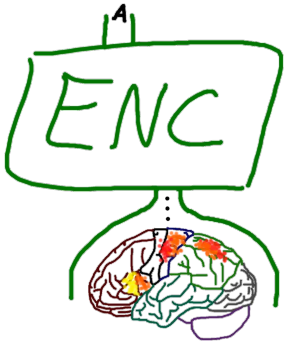


$$p(X_i|C = c_1) \stackrel{?}{\neq} p(X_i|C = c_2) \qquad X_i \not\perp C$$



$$p(X_i|C = c_1) \stackrel{?}{\neq} p(X_i|C = c_2) \qquad X_i \not\perp C$$

$$p(C|\vec{X}) \stackrel{?}{\neq} p(C|\vec{X} \setminus X_i)$$



$$p(X_i|C = c_1) \stackrel{?}{\neq} p(X_i|C = c_2) \qquad X_i \not\perp C$$

$$p(C|\vec{X}) \stackrel{?}{\neq} p(C|\vec{X} \setminus X_i) \qquad X_i \not\perp C|\vec{X} \setminus X_i$$

Causal interpretation rules (1)

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		
	√		
		×	
		√	
Response-based	×		
	√		
		×	
		√	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		
	✓		$X_i \not\perp R$
		×	
		✓	$X_i \leftarrow h \rightarrow R$
Response-based	×		$X_i \rightarrow R$
	✓		
		×	
		✓	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		
	√		
		×	
		√	
Response-based	×		
	√		
		×	
		√	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		
	✓		
		×	
		✓	
Response-based	×		
	✓		inconclusive
		×	
		✓	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		
	√		
		×	inconclusive
		√	inconclusive
Response-based	×		
	√		inconclusive
		×	inconclusive
		√	inconclusive

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	×		no effect of S
	√		effect of S
		×	inconclusive
		√	inconclusive
Response-based	×		no cause of R
	√		inconclusive
		×	inconclusive
		√	inconclusive

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	$\surd$	$\surd$	
	$\surd$	$\times$	
	$\times$	$\surd$	
	$\times$	$\times$	
Response-based	$\surd$	$\surd$	
	$\surd$	$\times$	
	$\times$	$\surd$	
	$\times$	$\times$	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	√	√	
	√	×	
	×	√	
	×	×	
Response-based	√	√	inconclusive
	√	×	
	×	√	
	×	×	

	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	✓	✓	$X_i \not\perp\!\!\!\perp S$ and $X_i \perp\!\!\!\perp S \vert \vec{X} \setminus X_i$
	✓	×	$S \dashrightarrow X_i$ indirectly
	×	✓	
	×	×	
Response-based	✓	✓	inconclusive
	✓	×	
	×	✓	
	×	×	

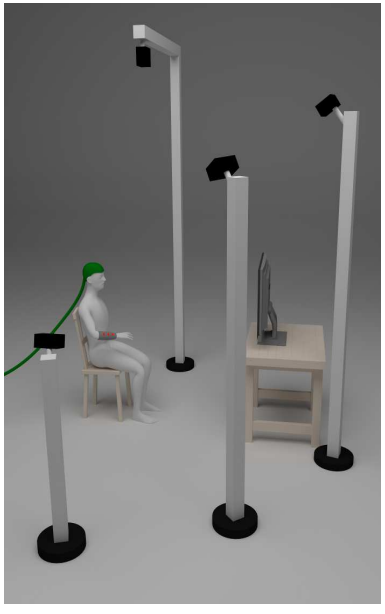
	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	√	√	
	√	×	
	×	√	
	×	×	
Response-based	√	√	inconclusive
	√	×	
	×	√	
	×	×	

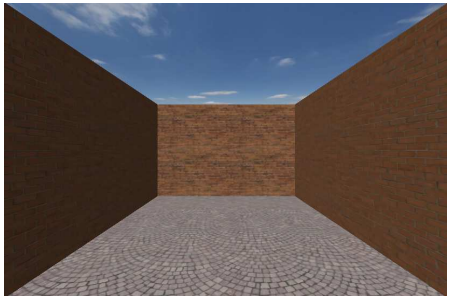
	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	√	√	indirect effect of S
	√	×	
	×	√	
	×	×	
Response-based	√	√	inconclusive
	√	×	
	×	√	
	×	×	

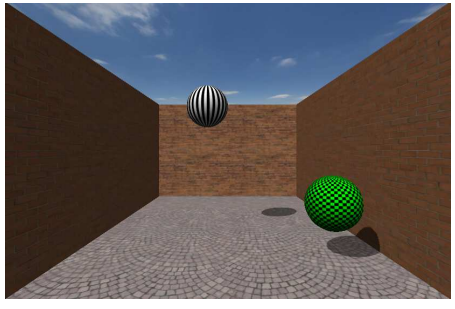
	Feature X_i relevant?		
	Encoding	Decoding	Causal interpretation
Stimulus-based	√	√	effect of S
	√	×	indirect effect of S
	×	√	provides context
	×	×	no effect of S
Response-based	√	√	inconclusive
	√	×	no direct cause of R
	×	√	provides context
	×	×	no cause of R

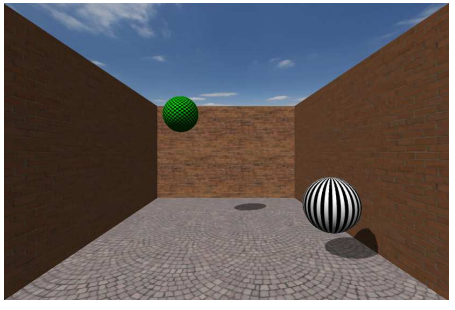
Empirical example

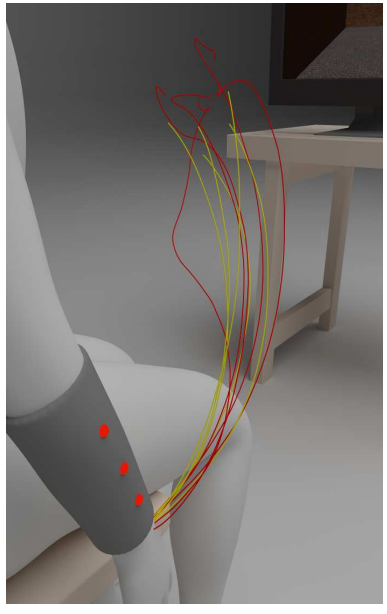
Visuo-motor learning task



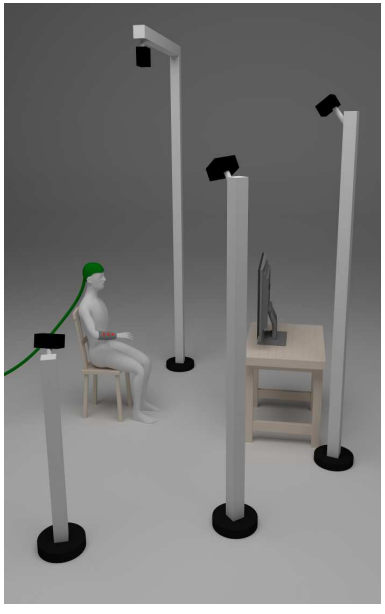


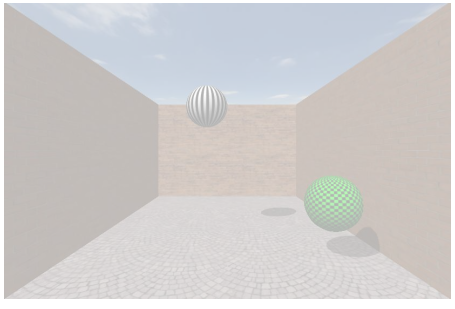


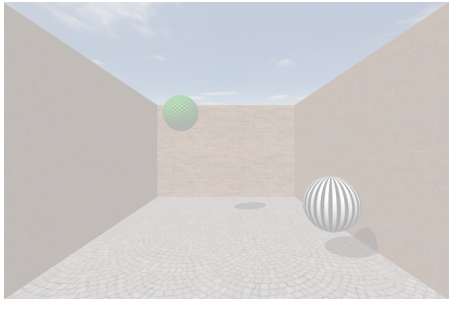


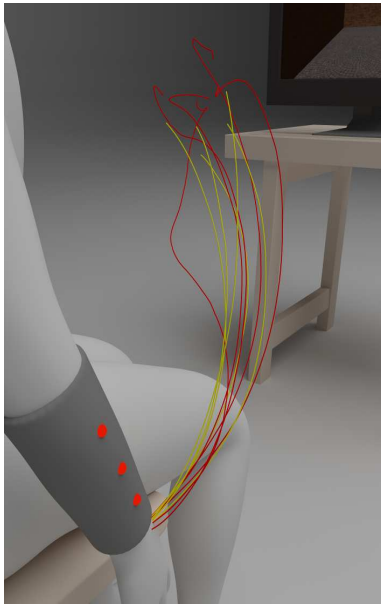


Visuo-motor learning task

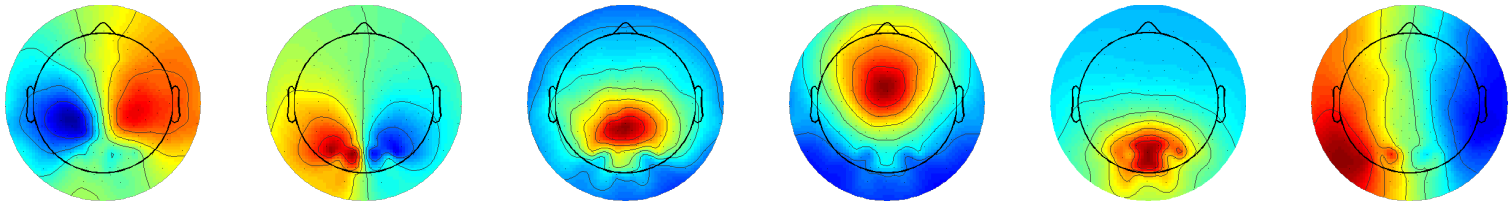




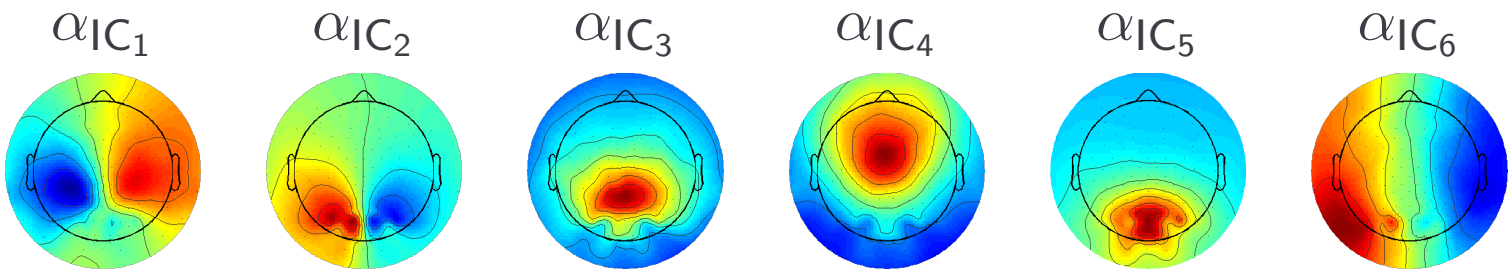


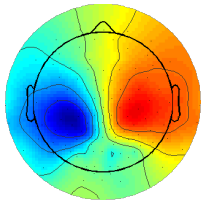
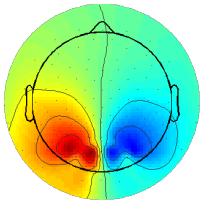
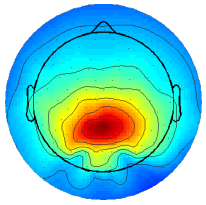
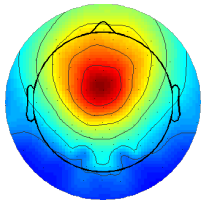
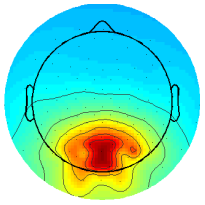
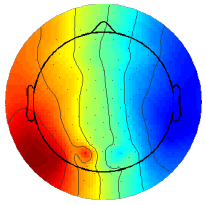


(Relevant) Features

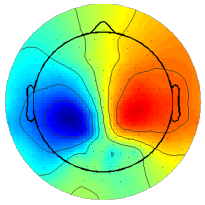
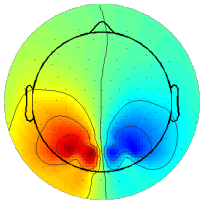
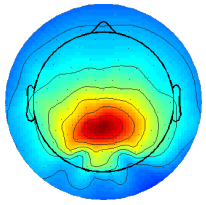
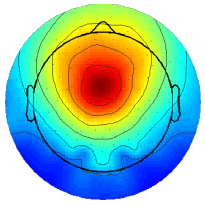
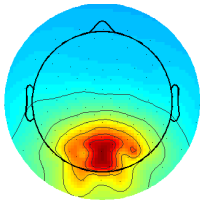
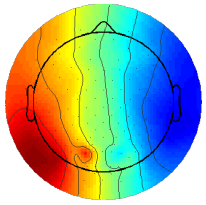


(Relevant) Features

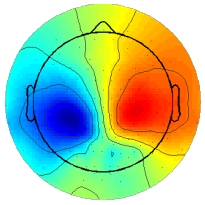
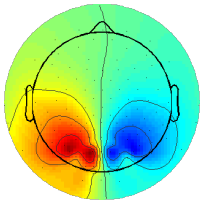
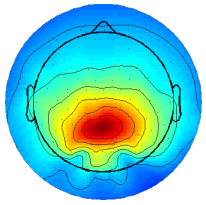
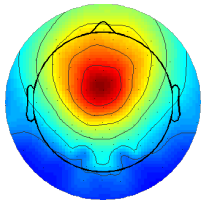
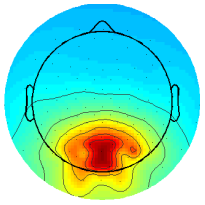
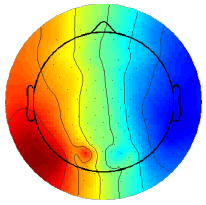


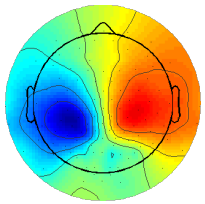
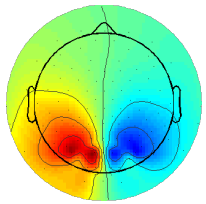
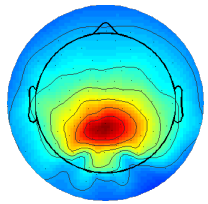
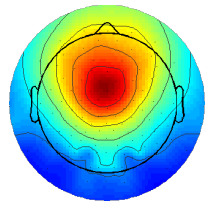
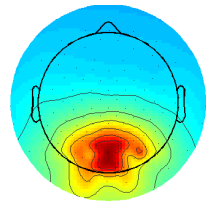
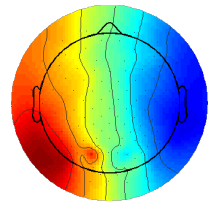
	α_{IC_1}	α_{IC_2}	α_{IC_3}	α_{IC_4}	α_{IC_5}	α_{IC_6}
p-values						
Encoding						
Decoding						

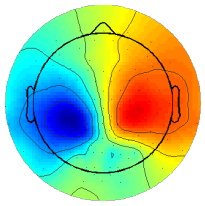
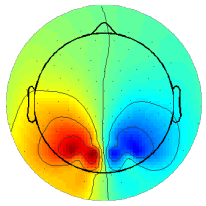
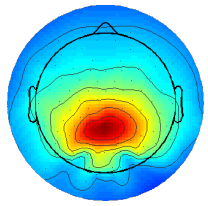
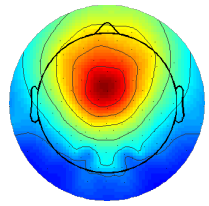
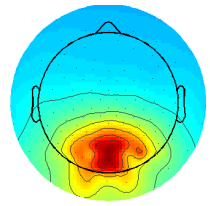
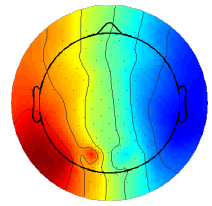
(Relevant) Features

p-values	α_{IC_1}	α_{IC_2}	α_{IC_3}	α_{IC_4}	α_{IC_5}	α_{IC_6}
						
Encoding	0	0	0	0	0	0
Decoding	0	0	0.50	0.34	0.79	0.13

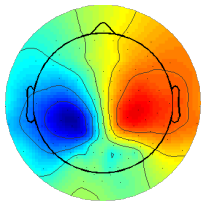
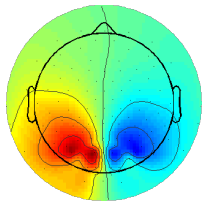
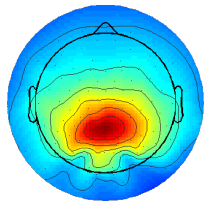
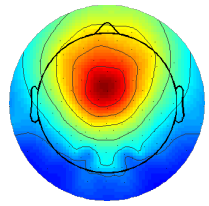
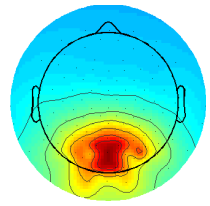
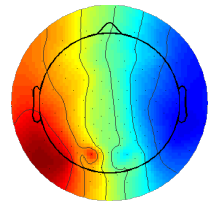
(Relevant) Features

p-values	α_{IC_1}	α_{IC_2}	α_{IC_3}	α_{IC_4}	α_{IC_5}	α_{IC_6}
						
Encoding	0	0	0	0	0	0
Decoding	0	0	0.50	0.34	0.79	0.13

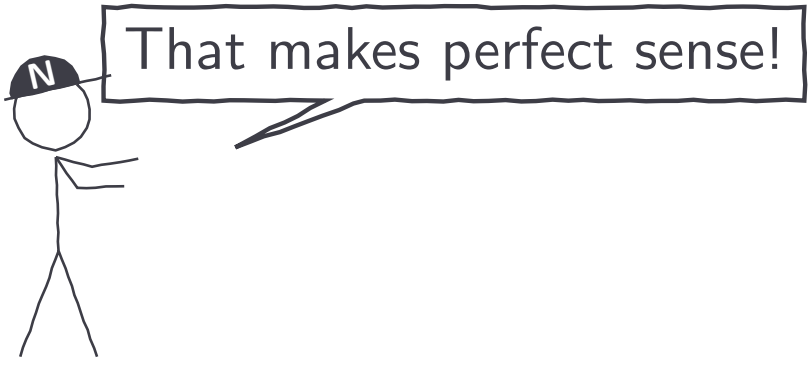
p-values	α_{IC_1} 	α_{IC_2} 	α_{IC_3} 	α_{IC_4} 	α_{IC_5} 	α_{IC_6} 
Encoding	0	0	0	0	0	0
Decoding	0	0	0.50	0.34	0.79	0.13

p-values	α_{IC_1} 	α_{IC_2} 	α_{IC_3} 	α_{IC_4} 	α_{IC_5} 	α_{IC_6} 
Encoding	0	0	0	0	0	0
Decoding	0	0	0.50	0.34	0.79	0.13

- instruction to plan a reaching movement is causal for all α_{IC_i}
- $\alpha_{IC_3}, \ldots, \alpha_{IC_6}$ are only indirect effects

p-values	α_{IC_1} 	α_{IC_2} 	α_{IC_3} 	α_{IC_4} 	α_{IC_5} 	α_{IC_6} 
Encoding	0	0	0	0	0	0
Decoding	0	0	0.50	0.34	0.79	0.13

- instruction to plan a reaching movement is causal for all α_{IC_i}
- $\alpha_{IC_3}, \ldots, \alpha_{IC_6}$ are only indirect effects



Wrap-up


N


C

feature relevance

feature relevance ⇝ (conditional) (in)dependence

feature relevance ⇝ (conditional) (in)dependence ⇝ causal structure

feature relevance ⇝ (conditional) (in)dependence ⇝ causal structure

- simple interpretation rules
- reinterpretation of previous results?
- resolve recently discussed issues

feature relevance ⇝ (conditional) (in)dependence ⇝ causal structure

- simple interpretation rules
- reinterpretation of previous results?
- resolve recently discussed issues

It's an interesting application!

Sebastian Weichwald, Timm Meyer, Ozan Özdenizci, Bernhard Schölkopf, Tonio Ball, Moritz Grosse-Wentrup:

- Causal interpretation rules for encoding and decoding models in neuroimaging. *NeuroImage*, 2015.
- Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014.