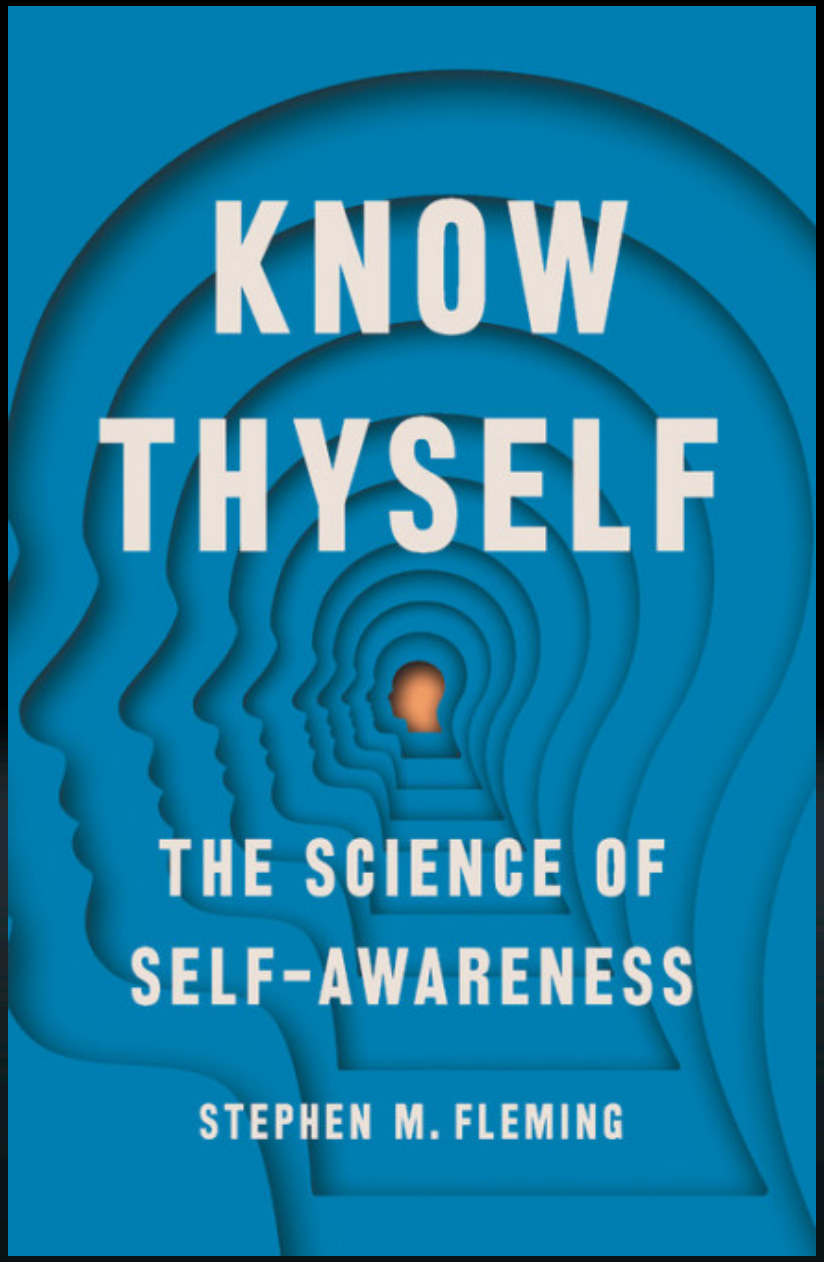


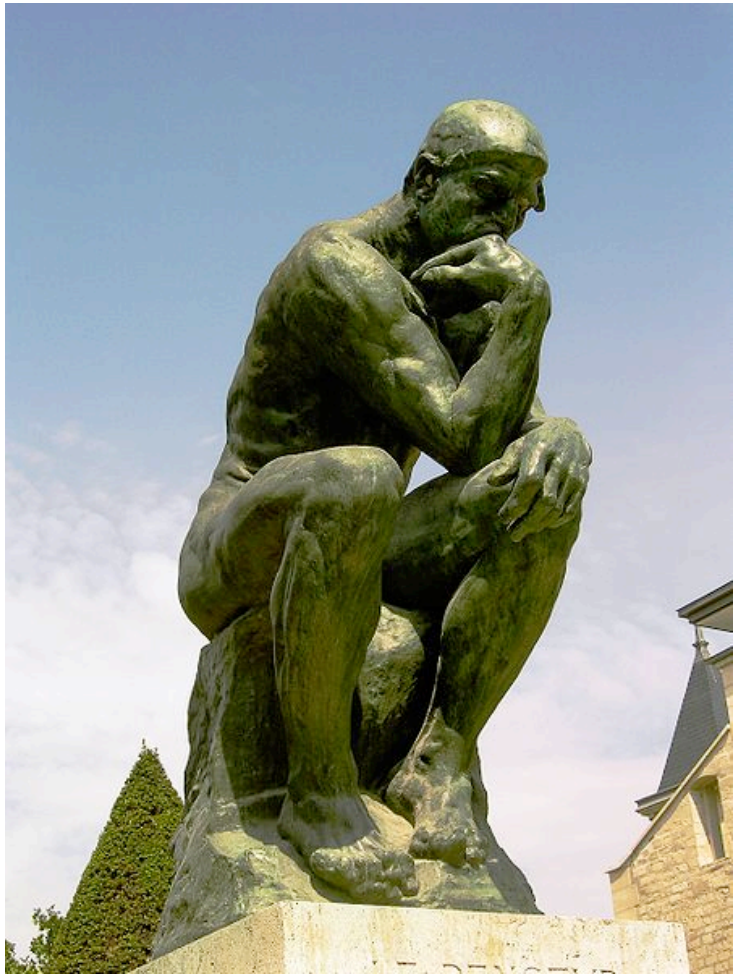
metacoglab.org/book

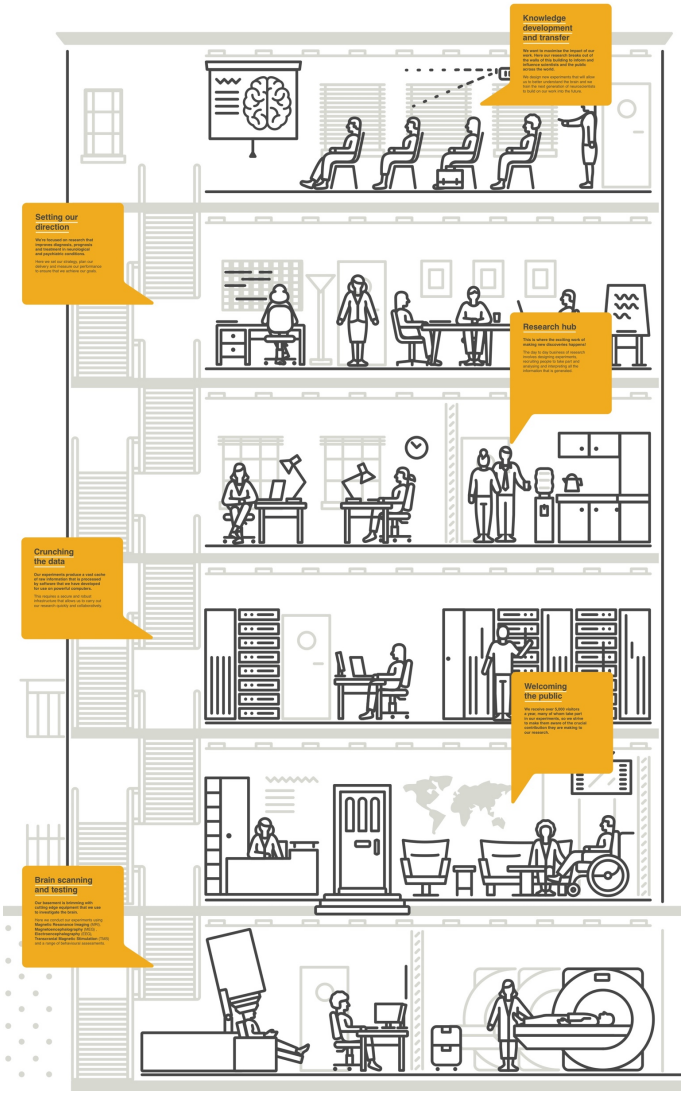
@smfleming

Being self-aware

"Being aware of being aware of being… In other words, if I not only know that I am, but also know that I know it, then I belong to the human species. All the rest follows - the glory of thought, poetry, a vision of the universe. In that respect the gap between ape and man is immeasurably greater than one between amoeba and ape."

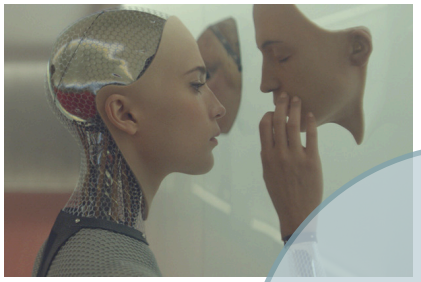
Vladimir Nabokov*, Strong Opinions*





wellcome centre human neuroimaging

Mental health



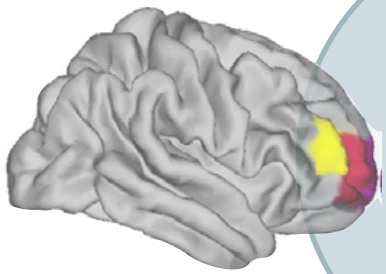
Consciousness

Decision-making

Vote Leave

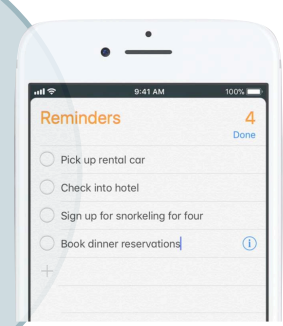
I'M IN
strongerin.co.uk

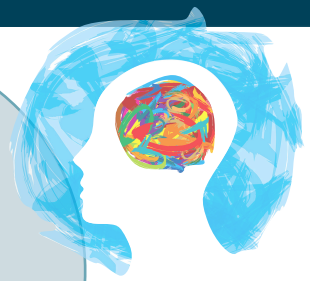
Metacognition of
- perception
- action
- memory



Prefrontal function

Cognitive offloading





Mental health

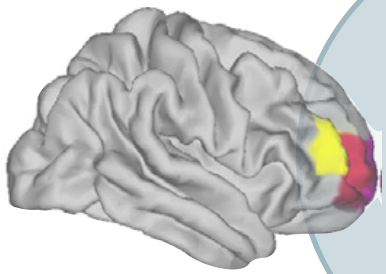
Consciousness

Metacognition of
- perception
- action
- memory

Decision-making

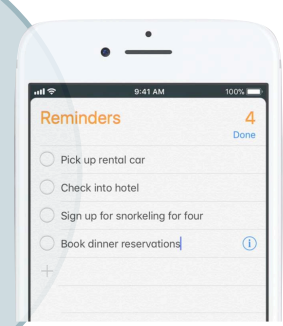
Vote Leave

I'M IN
strongerin.co.uk



Prefrontal function

Cognitive offloading



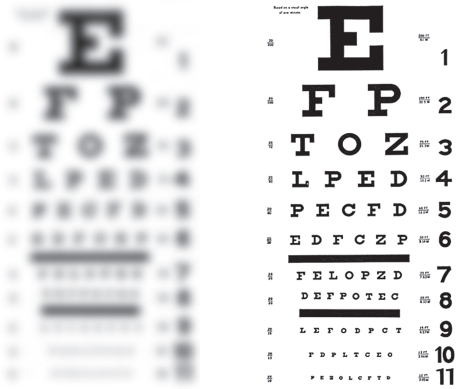
Metacognition - “thinking about thinking”

Meta-memory

Do I know this topic?

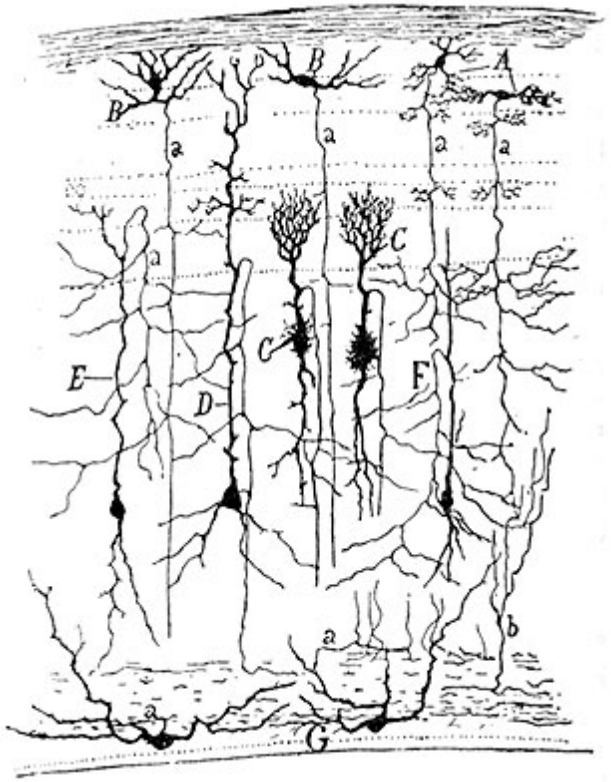
Meta-perception

Am I seeing things clearly?

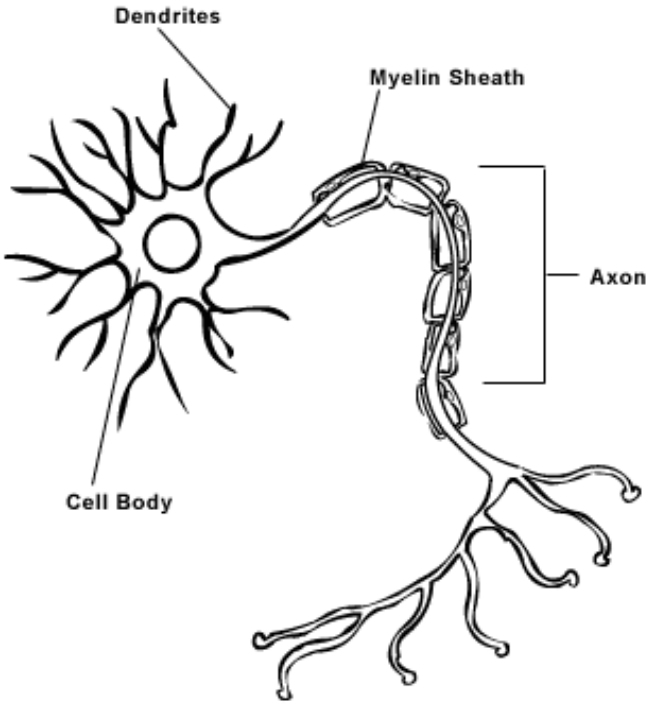


Google

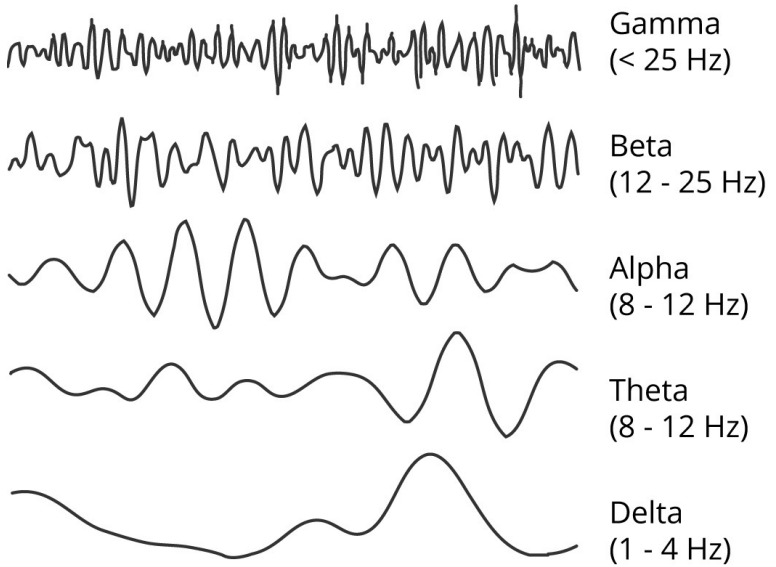
The raw materials



Santiago Ramon y Cajal,
1852 - 1934

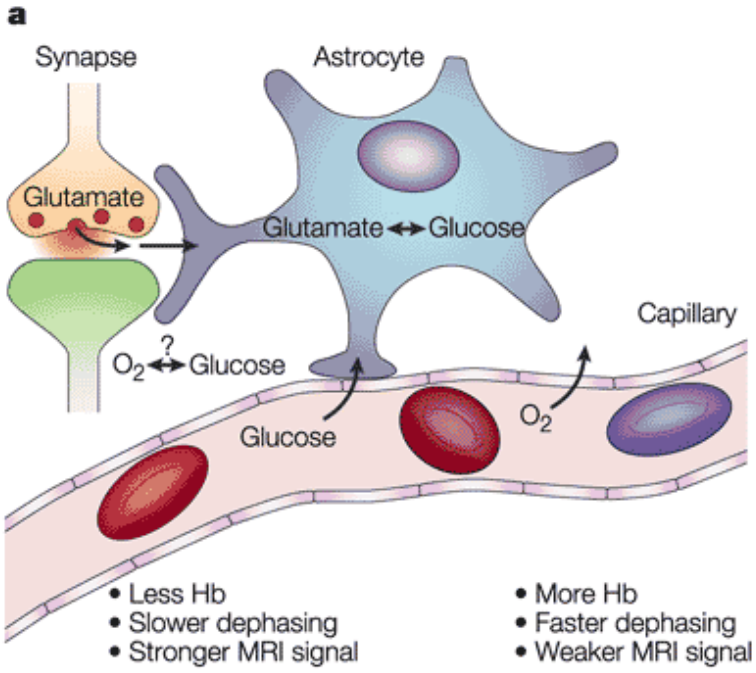


The biological basis of thought

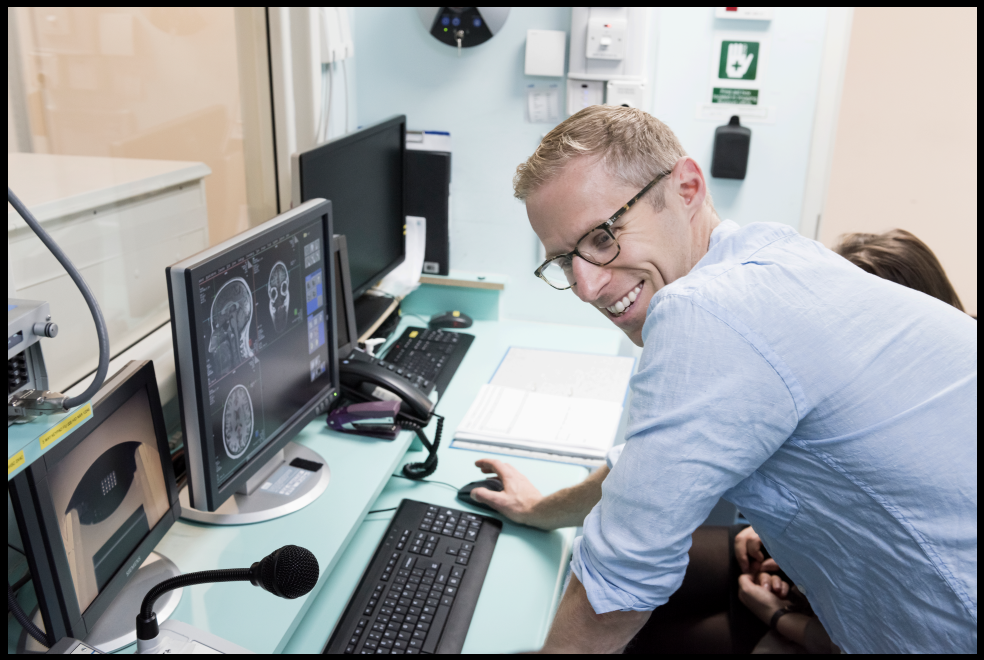


M/EEG – millisecond temporal resolution, poorer spatial resolution

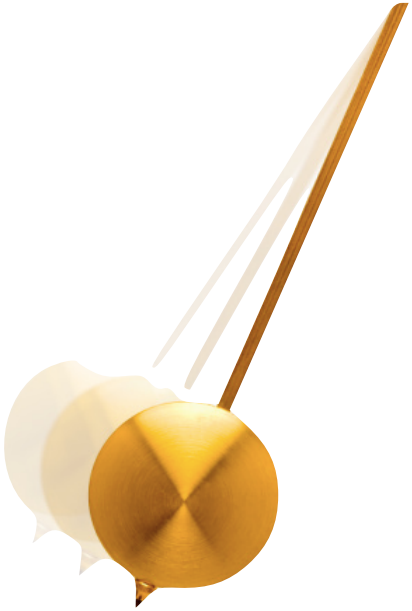
The biological basis of thought



fMRI – slow, high spatial resolution



How to measure metacognition



Introspection, self-report

Behaviorism / psychophysics

Ericsson & Simon (1984)

Nisbett & Wilson (1977)

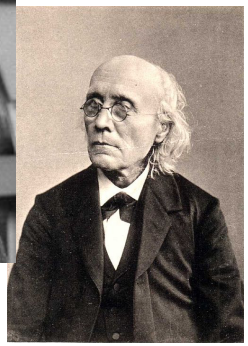
Metacognition (1980s-)

"skeptical introspectionists"

W. Wundt

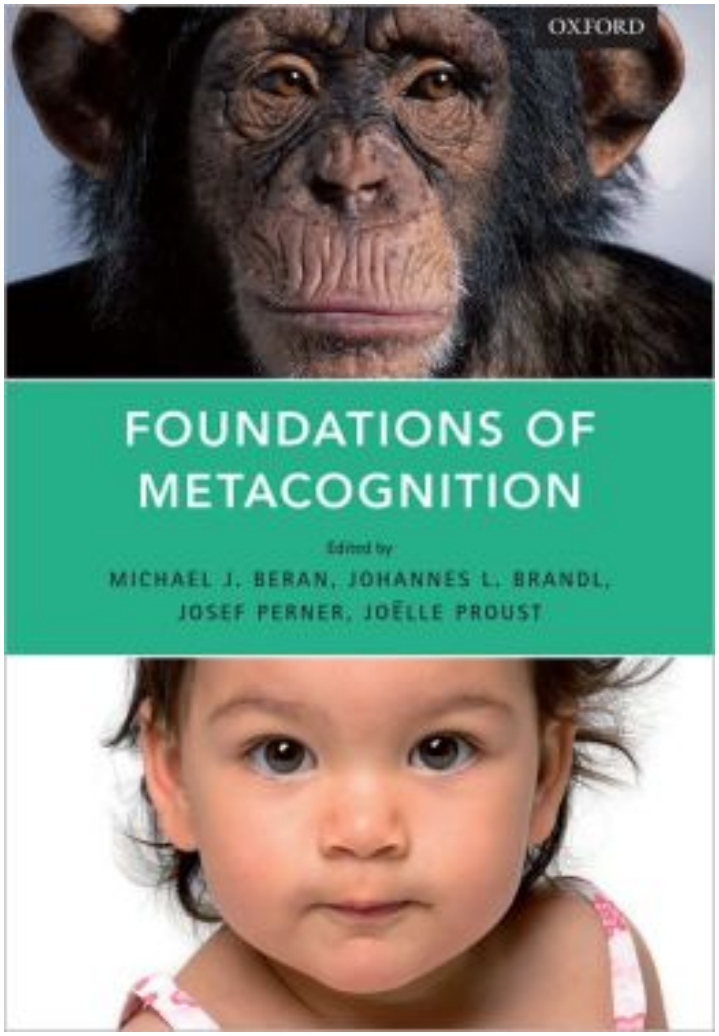
E. Titchener

B. F. Skinner



G. Fechner

A primer on measuring metacognition



BEHAVIOUR

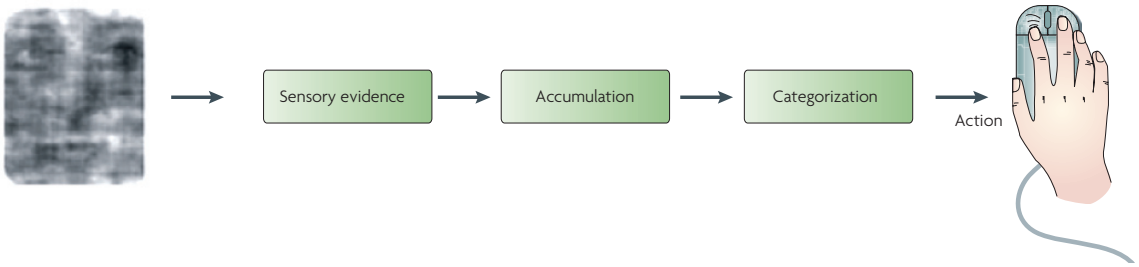
METACOGNITIVE JUDGMENT

E.g. answer to exam question; response in an experiment

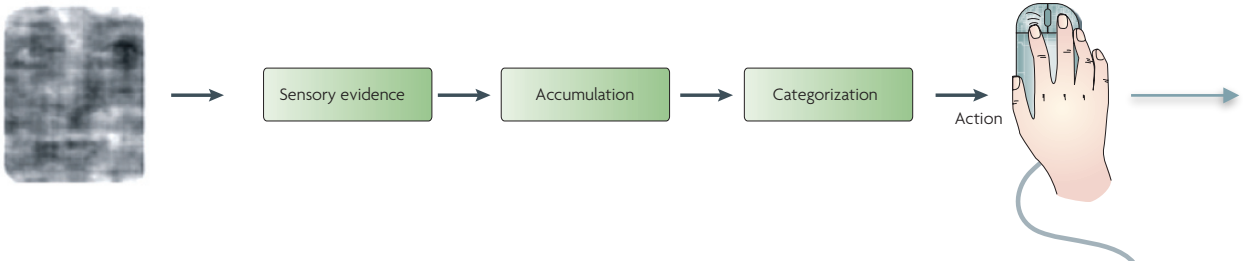
E.g. **confidence** in getting the answer right

Studying metacognition: Type 1 and Type 2 decisions

Type 1 decision:



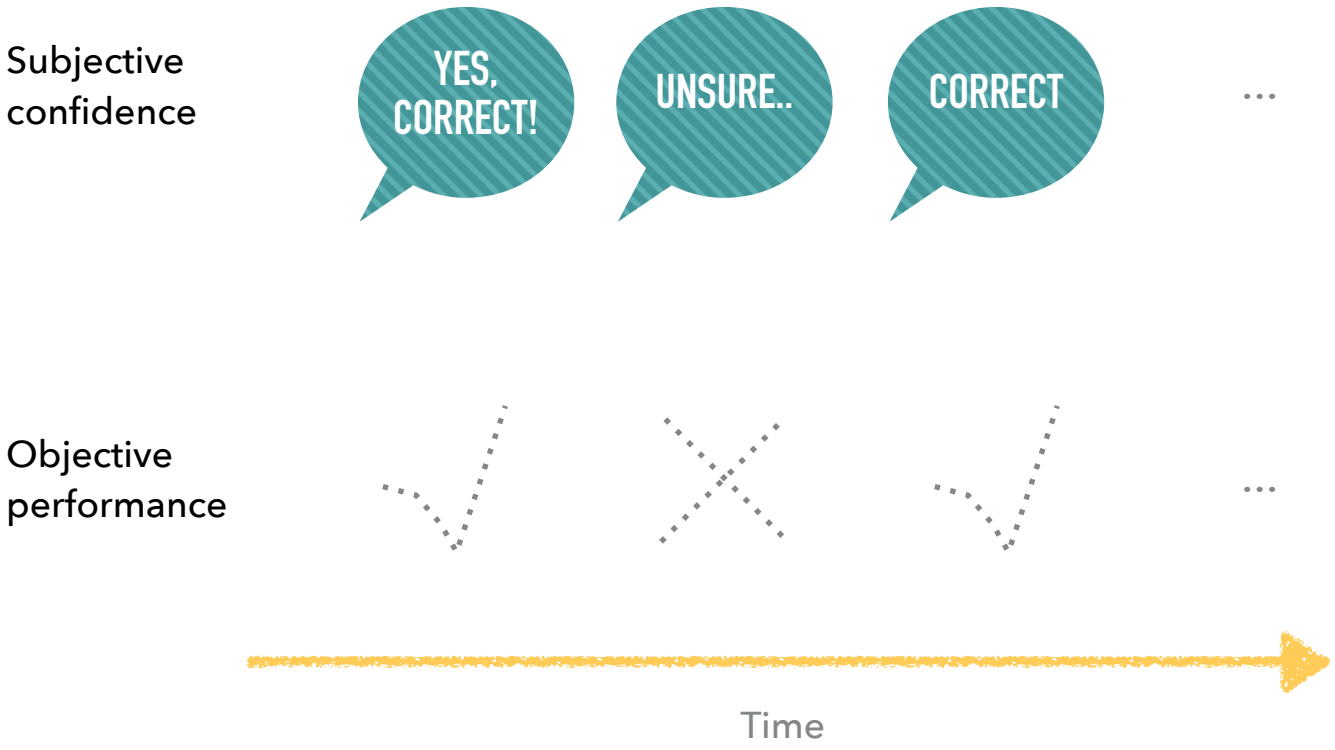
Type 2 decision - a decision "about" the type 1 decision:



How confident am I?
Did I make an error?

Not possible to quantify metacognition from a single judgment

Need multiple judgments over time, examine **statistical association** between behaviour and metacognitive judgments

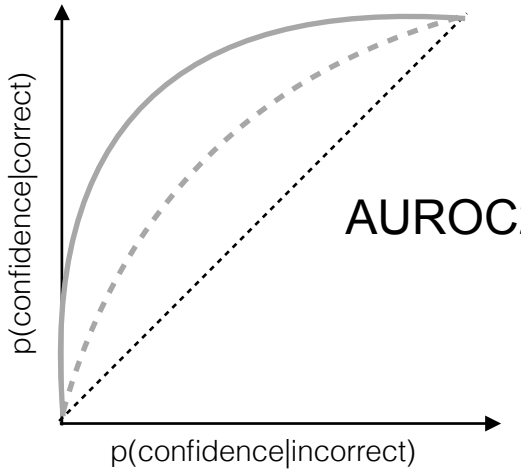


Quantifying metacognition - type 2 ROC analysis

Two Types of ROC Curves and Definitions of Parameters*

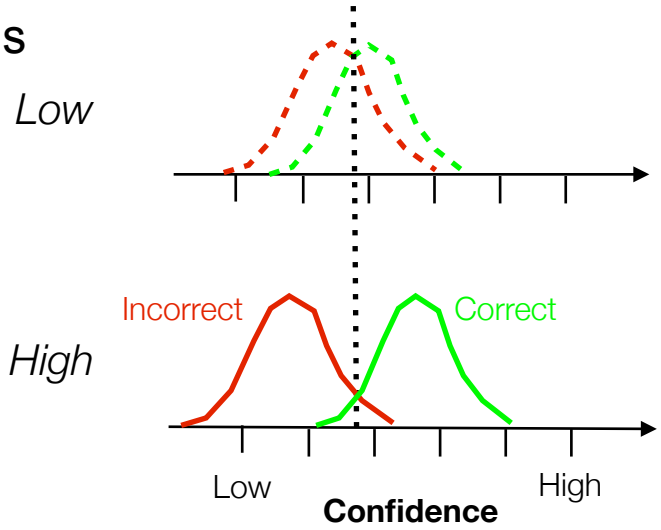
F. R. Clarke, T. G. Birdsall, and W. P. Tanner, Jr.
Electronic Defense Group, University of Michigan, Ann Arbor, Michigan
(Received February 26, 1959)

Type 2 ROC



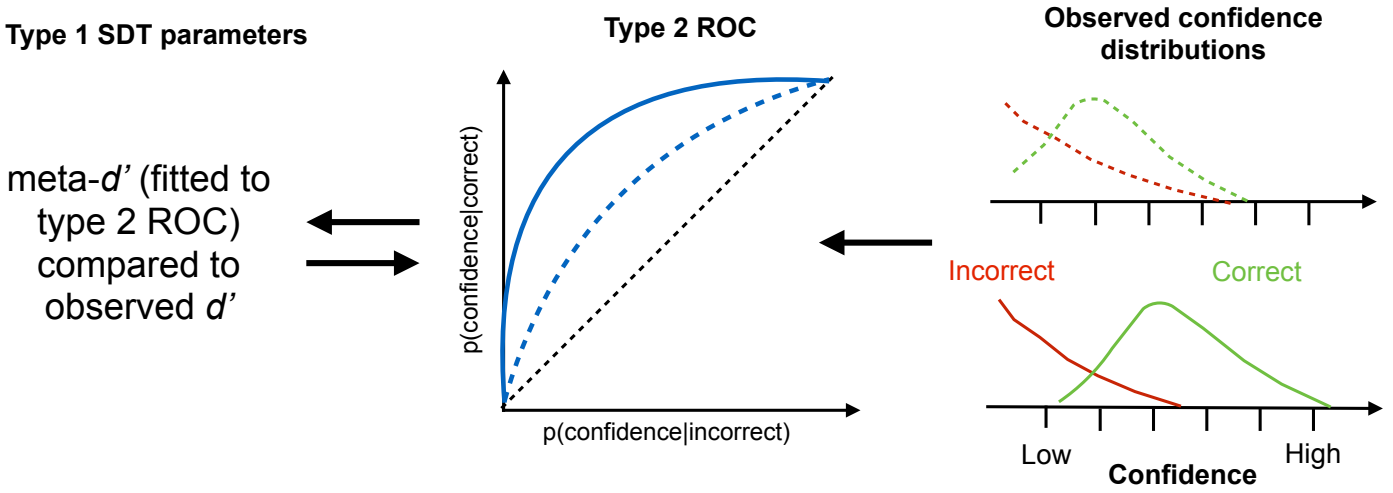
AUROC2

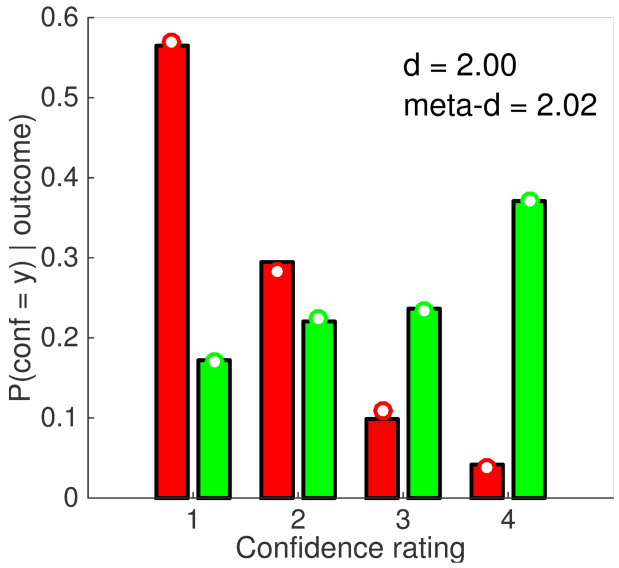
free from distributional assumptions

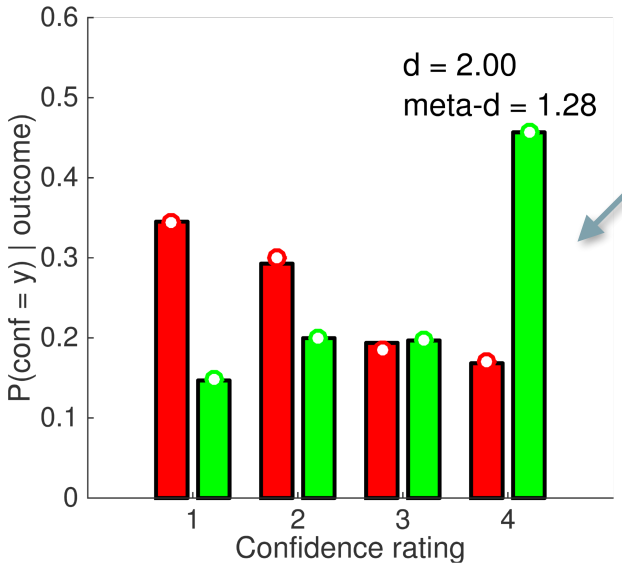


Clarke et al. (1959) *J Acoust Soc Am*; Galvin et al. (2003) *Psych Bull Rev*

Generative model for metacognition - meta-d'







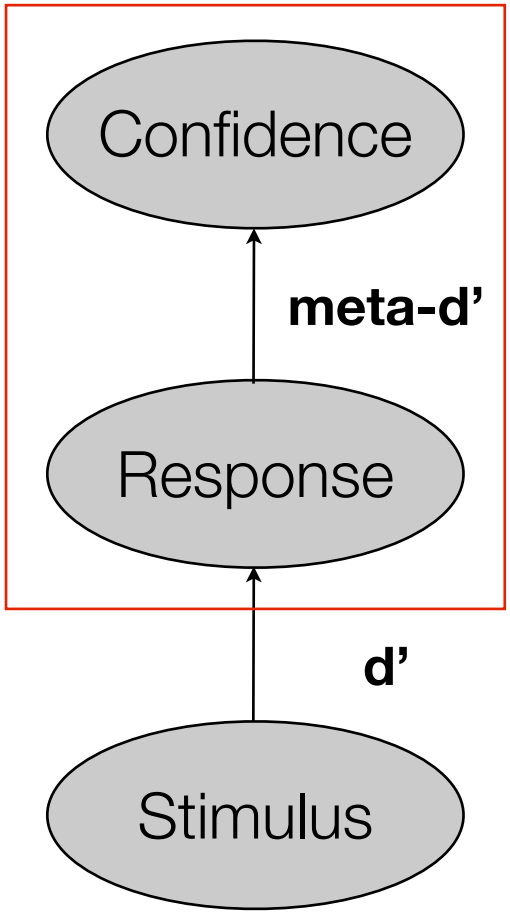
Gaussian noise added to confidence ratings

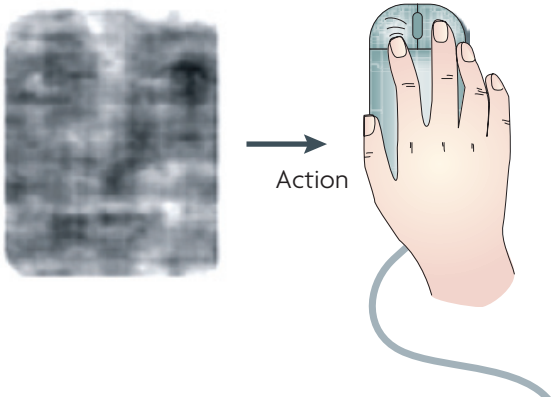
Maniscalco & Lau (2012) *Consciousness & Cognition*; Fleming (2017) *Neuroscience of Consciousness*

Type 1 and Type 2 sensitivity

Metacognitive (Type 2) sensitivity

Type 1 sensitivity



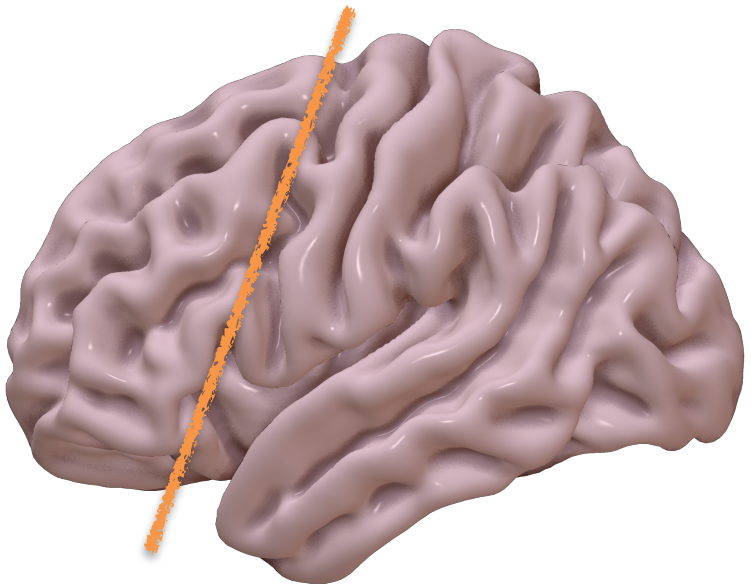


meta-d'/d' = metacognitive efficiency

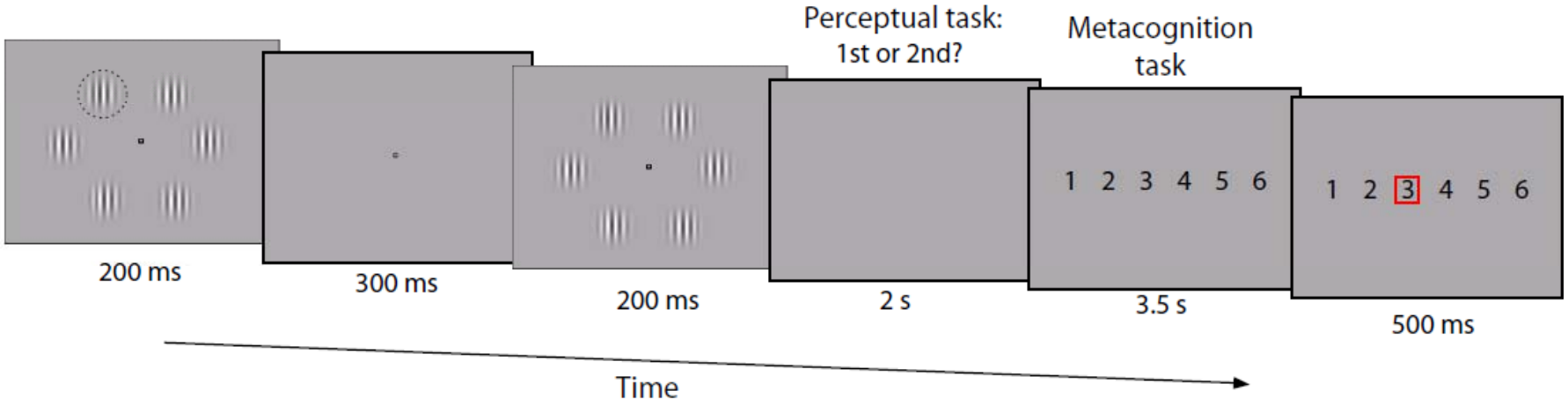
Maniscalco & Lau (2012) *Consciousness & Cognition*; Fleming (2017) *Neuroscience of Consciousness*

Individual differences in metacognition

Key focus on frontal lobe; damage often affects self-awareness / insight



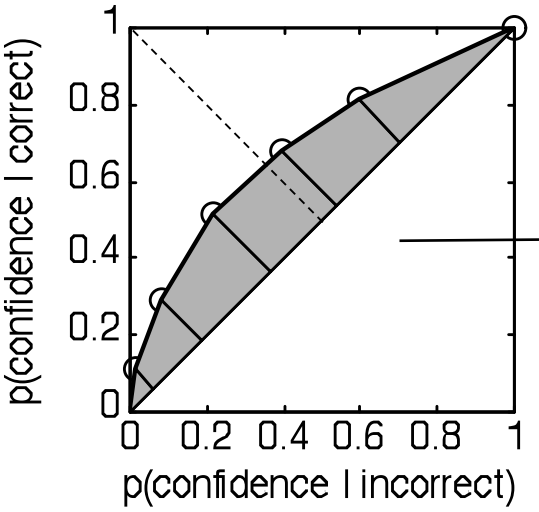
Isolating metacognition from performance

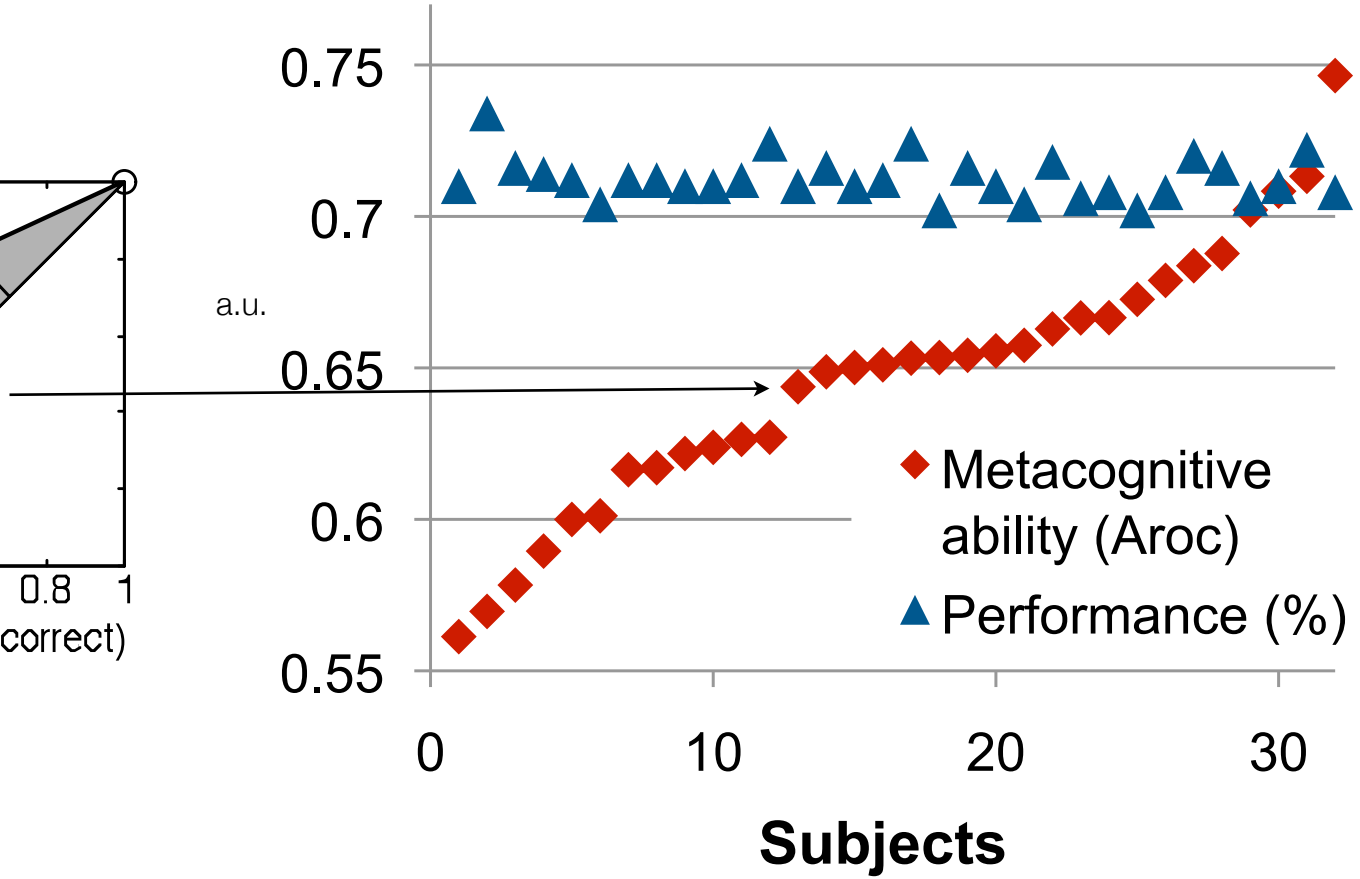


Performance titrated using a 2-down 1-up staircase
32 participants
600 trials per participant

Fleming et al. (2010) *Science*

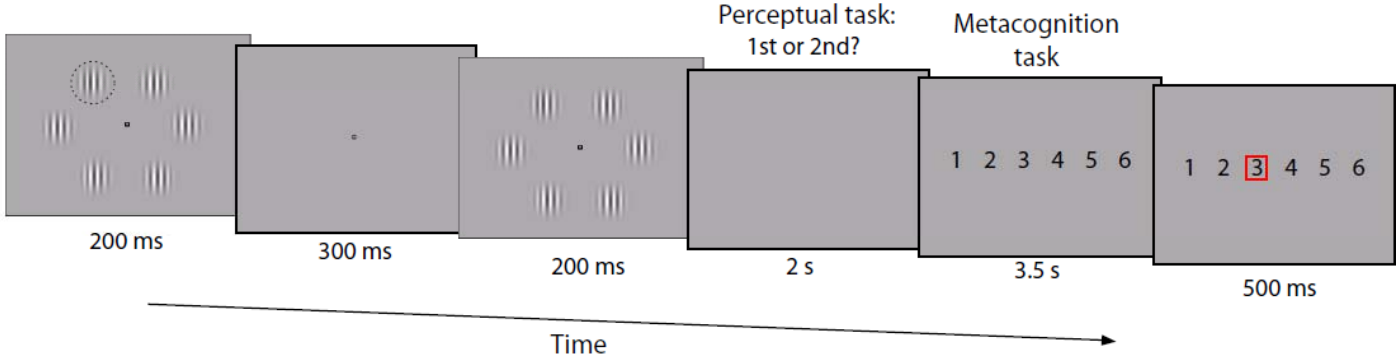
Isolating metacognition from performance

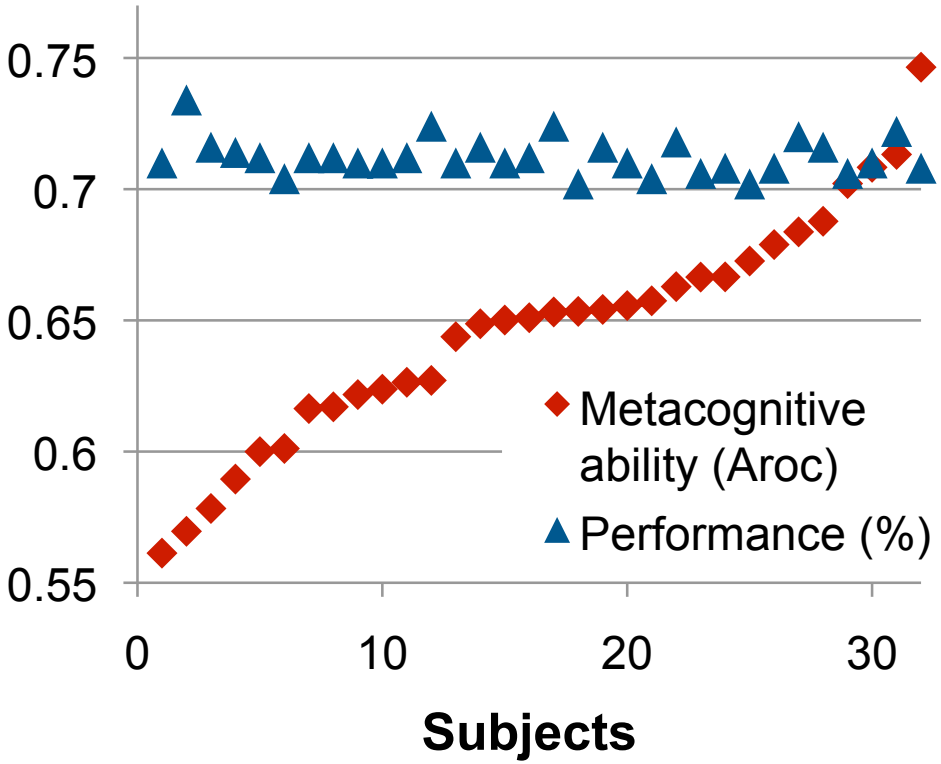


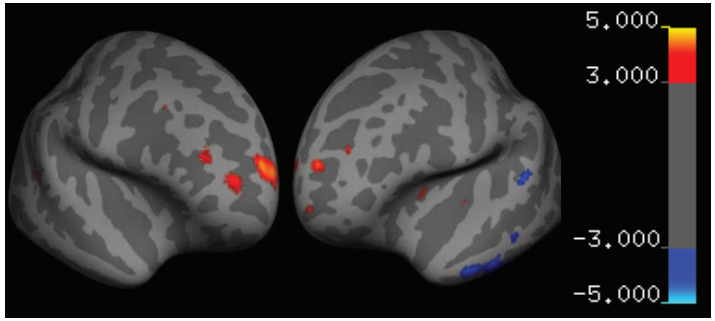


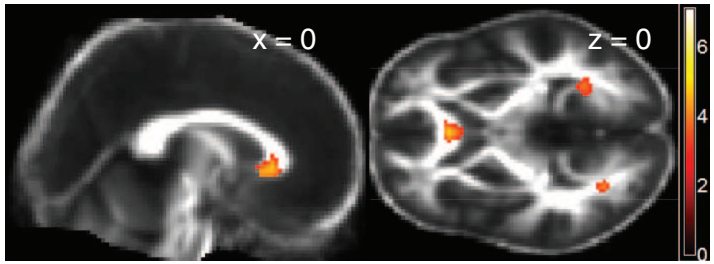
Fleming et al. (2010) *Science*

Isolating metacognition from performance



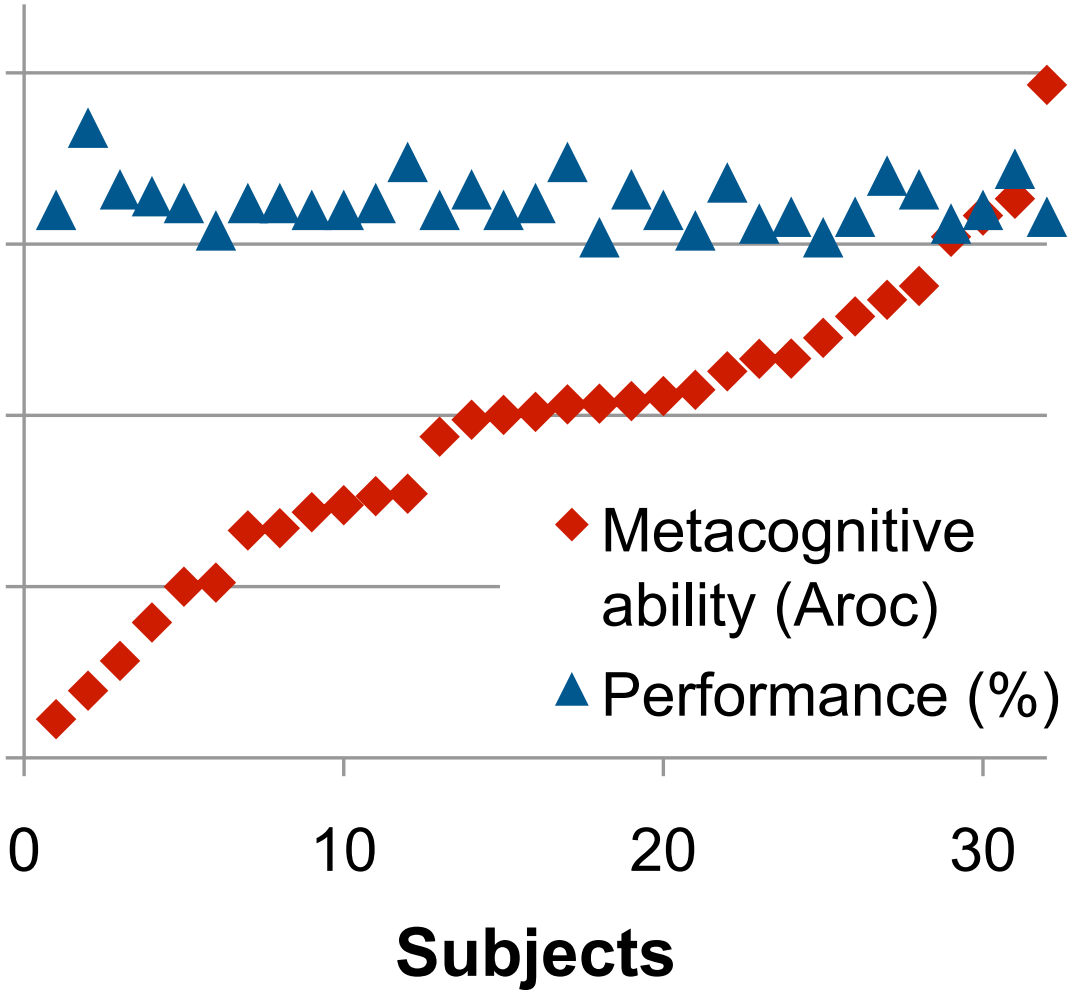


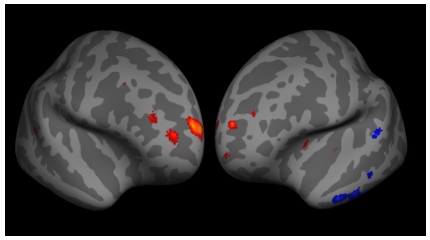




Fleming et al. (2010) *Science*

Individual variation in metacognition





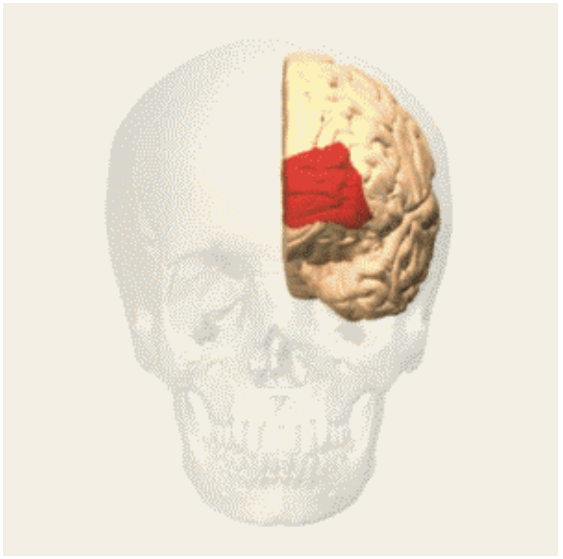
- Metacognitive sensitivity varies considerably across individuals
- Unrelated to gender or general cognitive ability (IQ)

Fleming et al. (2010) *Science*

Metacognitive sensitivity and aPFC

Relating Introspective Accuracy to Individual Differences in Brain Structure

Stephen M. Fleming,[1]*† Rimona S. Weil,[1,2]* Zoltan Nagy,[1] Raymond J. Dolan,[1] Geraint Rees[1,2]

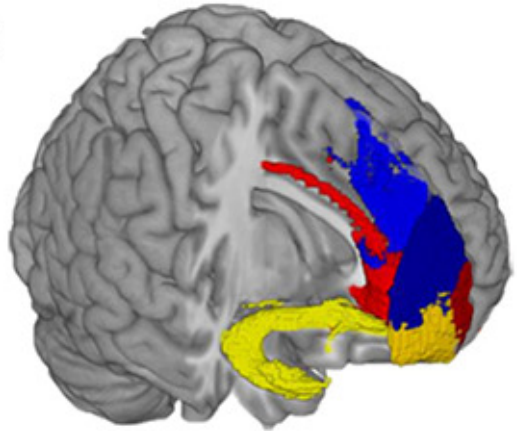


Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance

Osamu Yokoyama[a,b,c], Naoki Miura[c,d], Jobu Watanabe[c,d,e], Atsushi Takemoto[b,c], Shinya Uchida[d,f], Motoaki Sugiura[g], Kaoru Horie[e,h], Shigeru Sato[e,h], Ryuta Kawashima[c,d,e,f], Katsuki Nakamura[b,c,*]

Anatomical Coupling between Distinct Metacognitive Systems for Memory and Visual Perception

Li Yan McCurdy,[1] Brian Maniscalco,[1] Janet Metcalfe,[1] Ka Yuet Liu,[2] Floris P. de Lange,[3] and Hakwan Lau[1,3]
[1]Department of Psychology, Columbia University, New York, New York 10027, [2]Department of Sociology, University of California, Los Angeles, Los Angeles, California 90095, and [3]Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, 6500 HE Nijmegen, The Netherlands



Metacognitive ability correlates with hippocampal and prefrontal microstructure

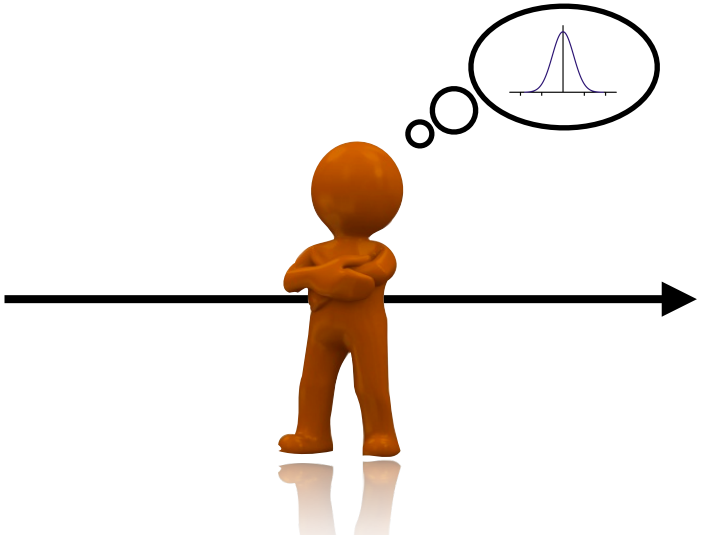
Micah Allen[a,b,*], James C. Glen[a], Daniel Müllensiefen[c], Dietrich Samuel Schwarzkopf[a,d], Francesca Fardo[a,e,f], Darya Frank[g], Martina F. Callaghan[b], Geraint Rees[a,b]

Medial and Lateral Networks in Anterior Prefrontal Cortex Support Metacognitive Ability for Memory and Perception

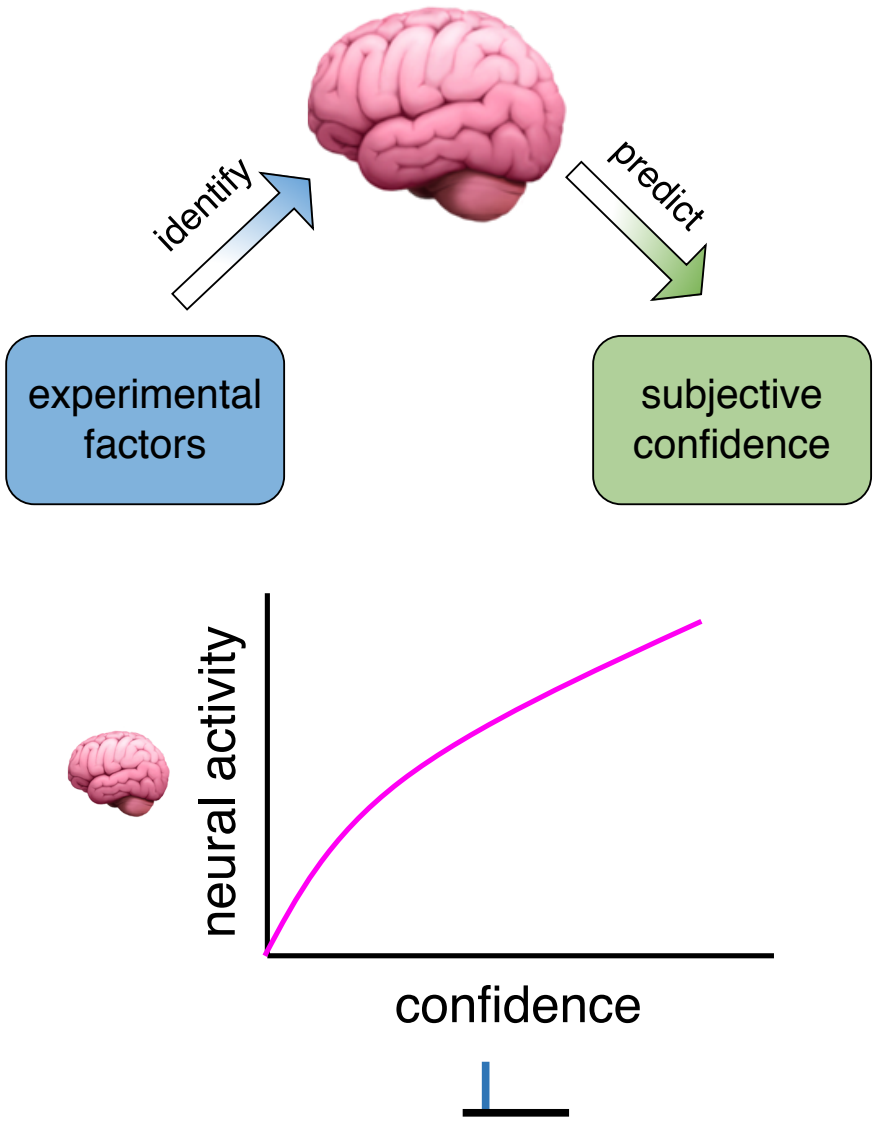
Benjamin Baird,[1] Jonathan Smallwood,[2] Krzysztof J. Gorgolewski,[3] and Daniel S. Margulies[3]
[1]Department of Psychological and Brain Sciences, University of California, Santa Barbara, California 93106, [2]Department of Psychology, University of York, North Yorkshire YO10 5DD, United Kingdom, and [3]Max Planck Research Group: Neuroanatomy & Connectivity, Max Planck Institute for Human Cognitive and Brain Sciences, 04103, Leipzig, Germany

Individual differences in metacognition

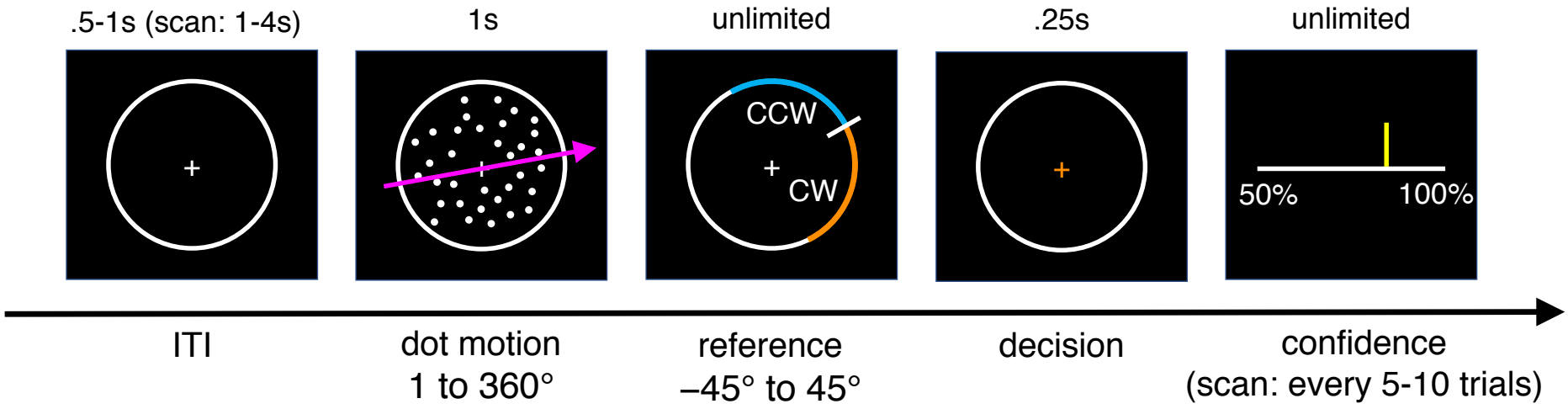
Within-subject construction of confidence



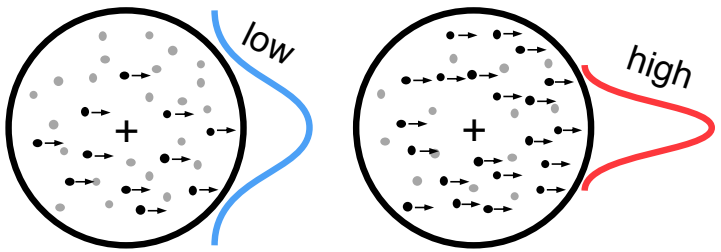
Neural encoding of confidence



Controlling influences on confidence

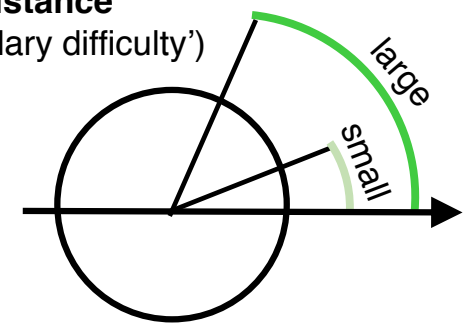


coherence
('sensory reliability')



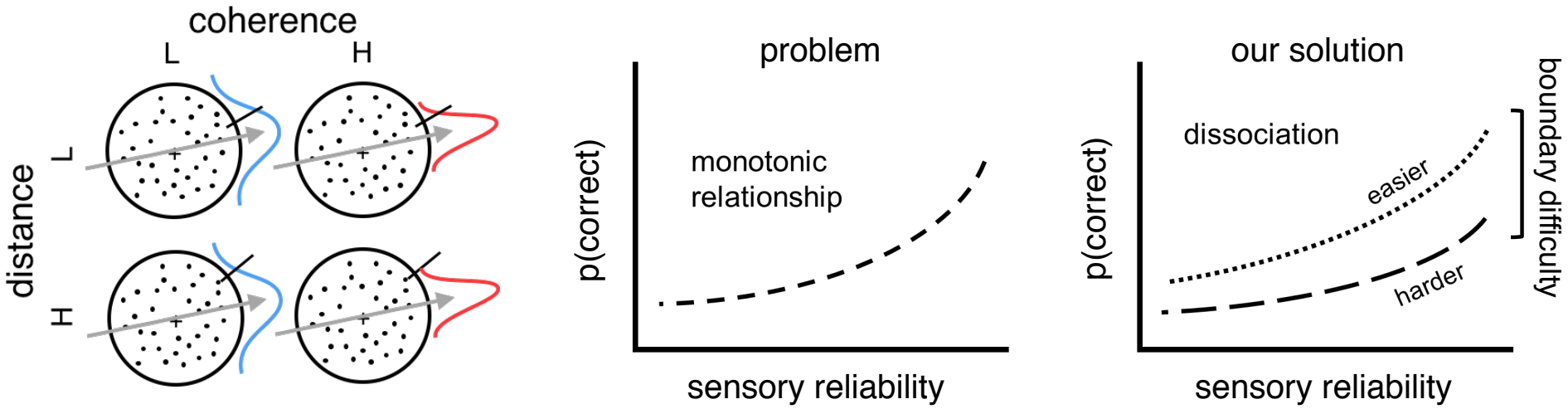
x

distance
('boundary difficulty')



Bang & Fleming (2018) *PNAS*

Distilling confidence from sensory reliability



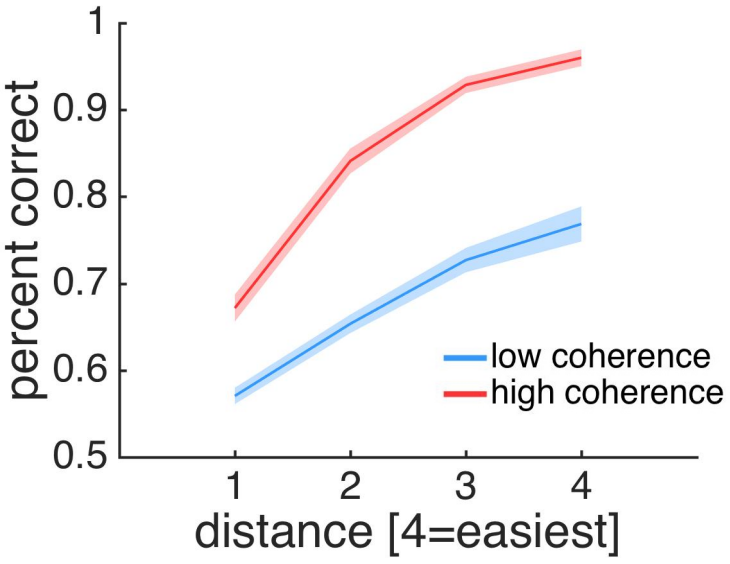
Factorial design decouples sensory reliability and confidence

Bang & Fleming (2018) *PNAS*

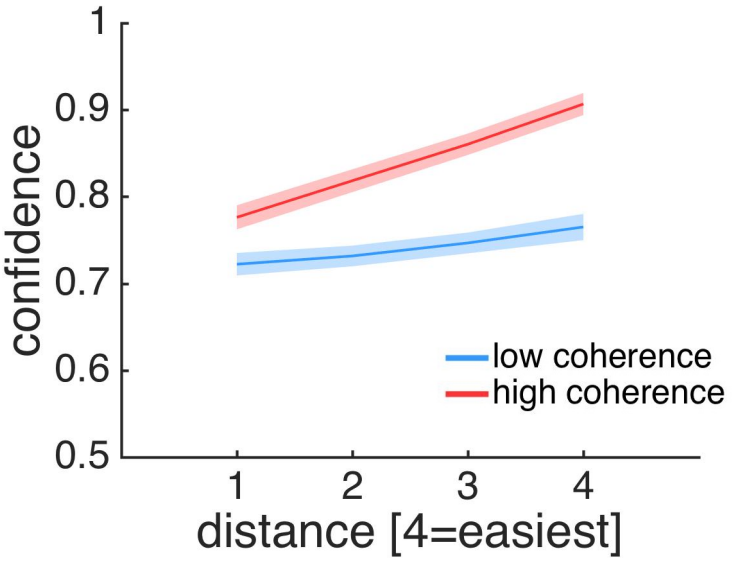
Distilling confidence from sensory reliability

Behavioural results

successfully introduced multiple influences on p(correct)

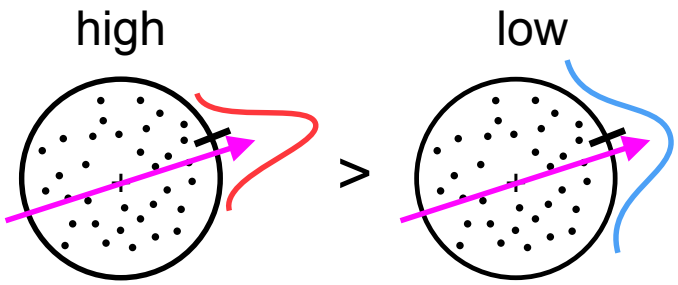


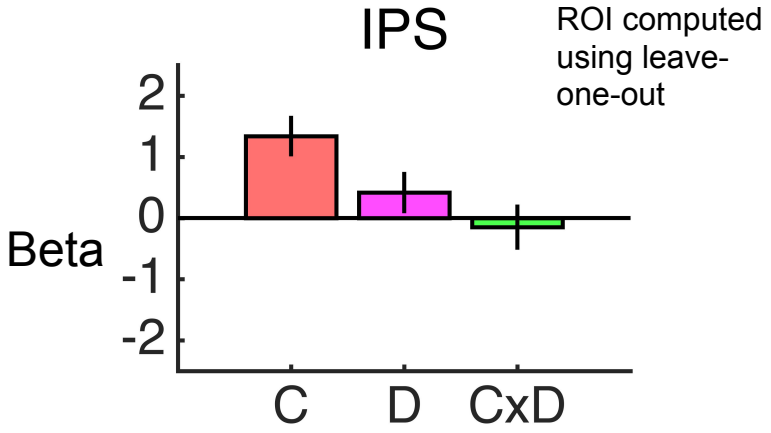
subjective reports reflect these influences

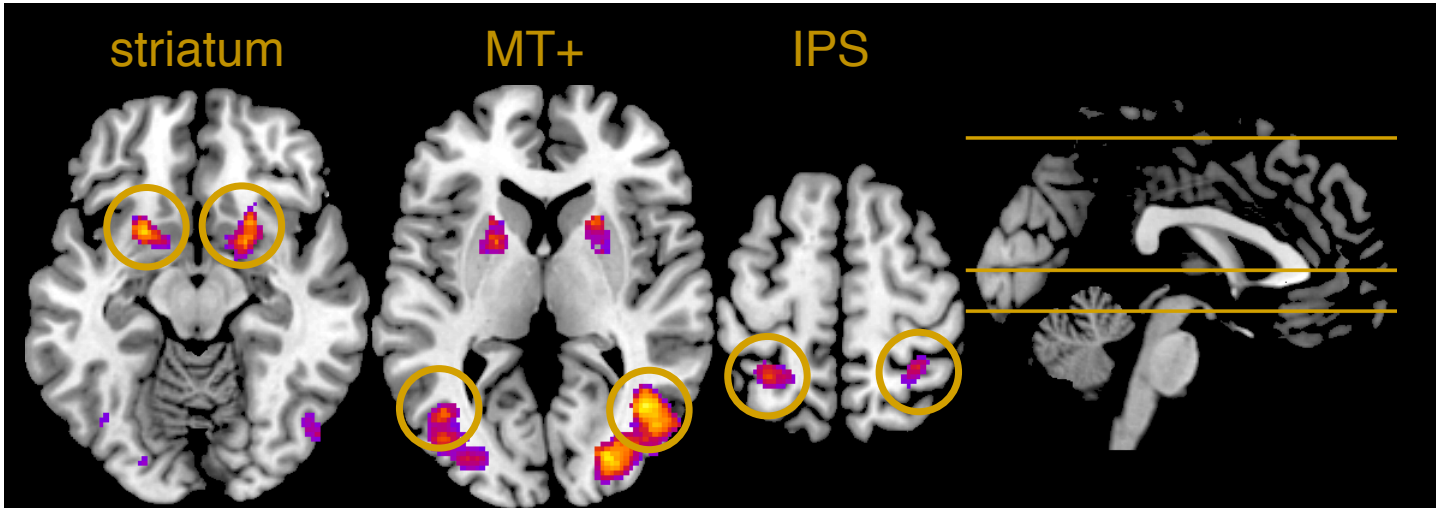


Bang & Fleming (2018) *PNAS*

Components of confidence - sensory reliability





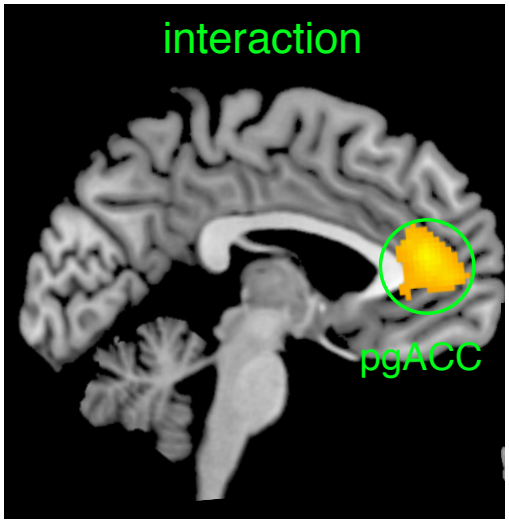


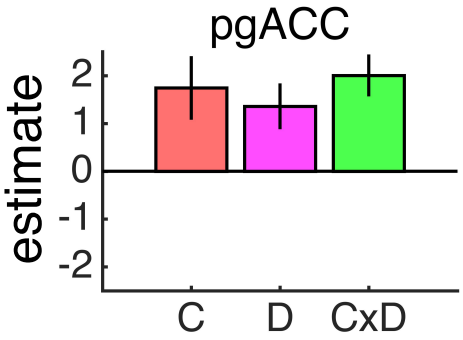
cluster-defining threshold: $p < .001$; $pFWE < .05$ corrected; correct trials

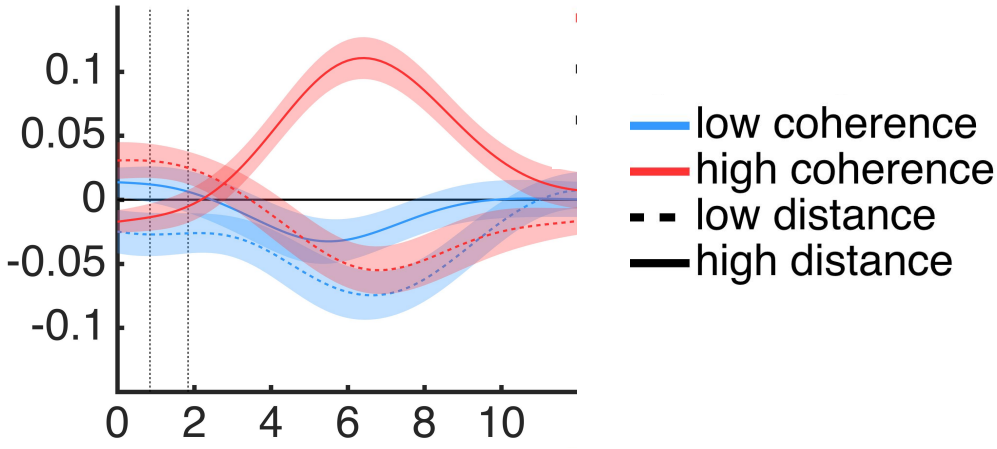
Bang & Fleming (2018) *PNAS*

Distinct neural signature of confidence in PFC

Medial PFC tracks certainty, distance and their interaction - experimental markers of confidence



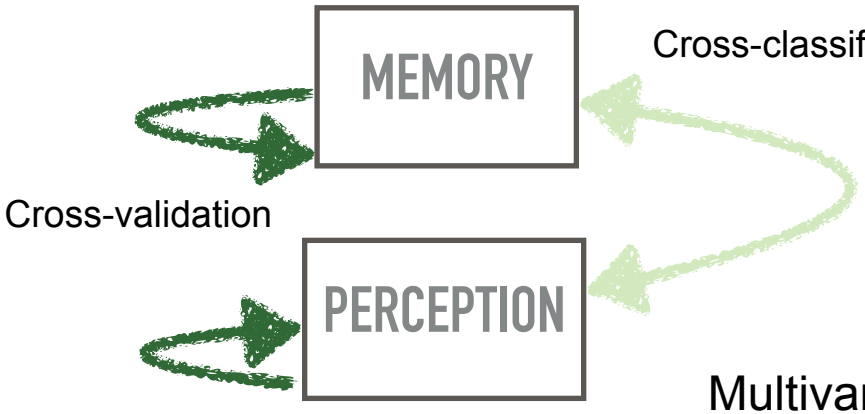


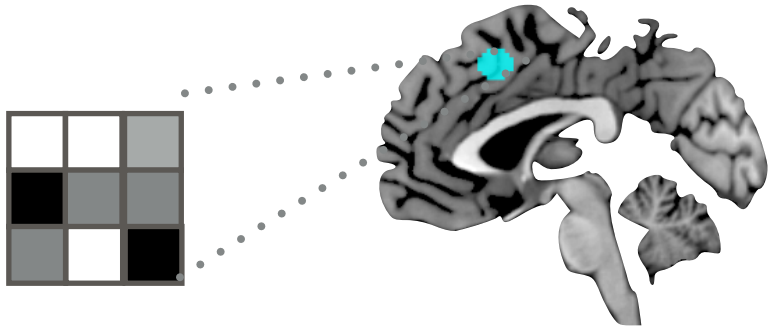


Findings suggest a **dissociation** between neural populations tracking **sensory certainty** (parietal cortex) and those tracking **decision confidence** (PFC)

Bang & Fleming (2018) *PNAS*

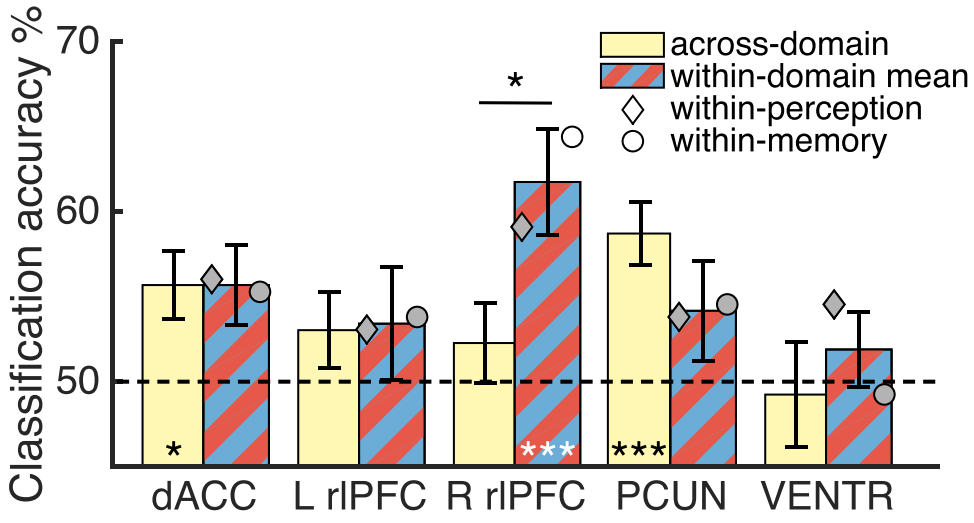
Shared signals for confidence across tasks

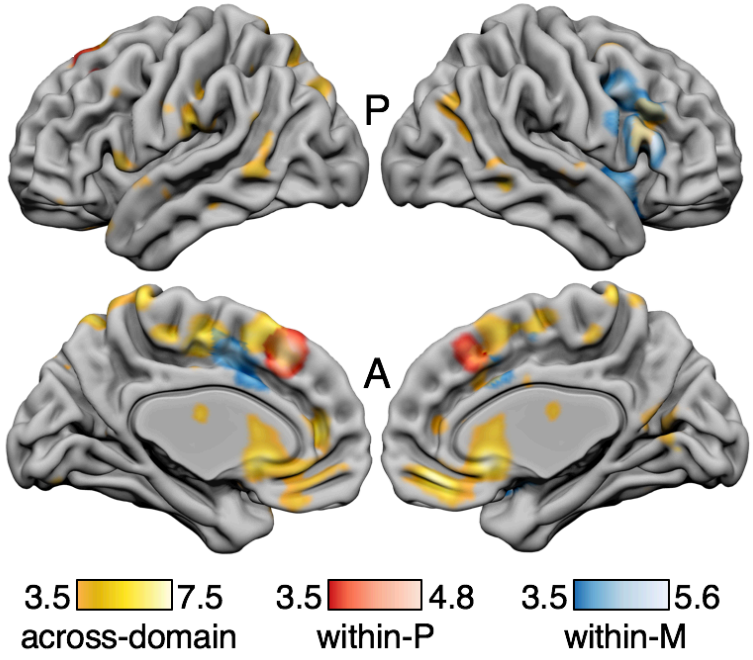




Multivariate pattern analysis (MVPA)

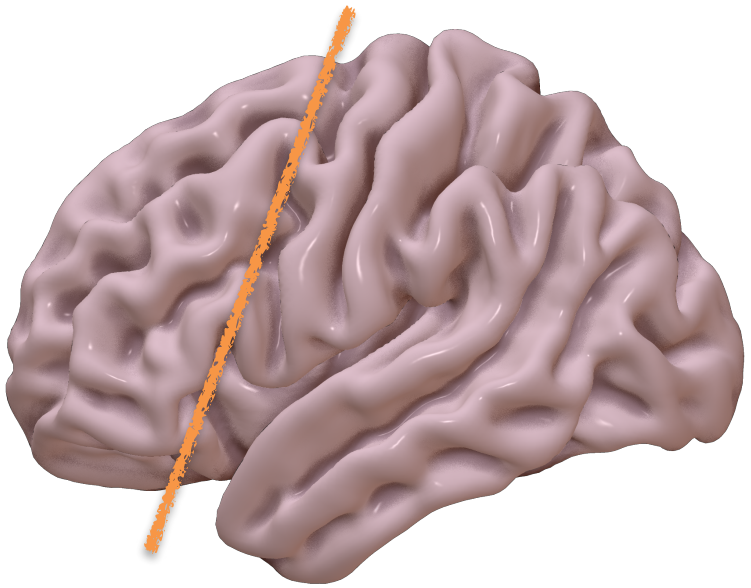
Confidence-related activity patterns

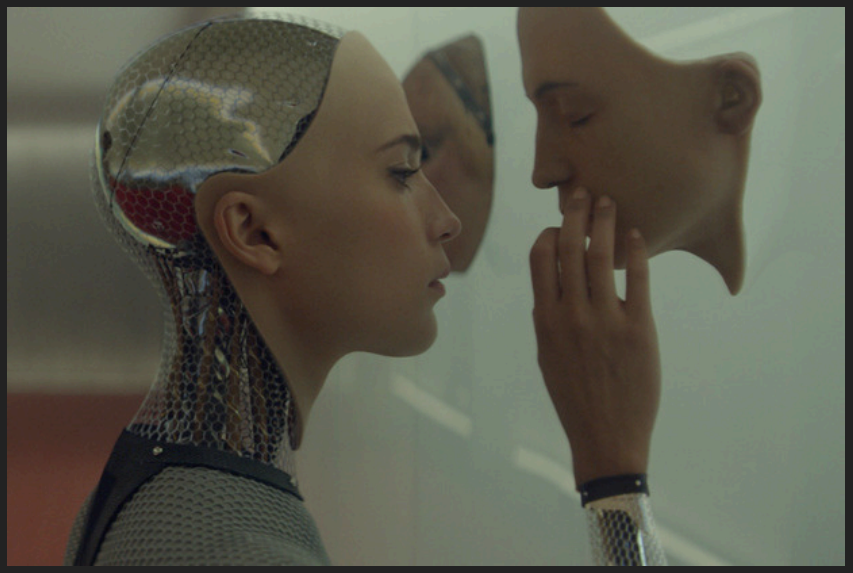




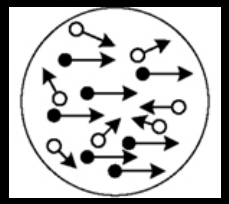
Morales, Lau & Fleming (2018) *J Neuro*

Recruitment of prefrontal cortex in service of metarepresentations can explain loss of self-awareness in brain damage/disease





Computing confidence: basic ingredients



World state (*d*)

L or R

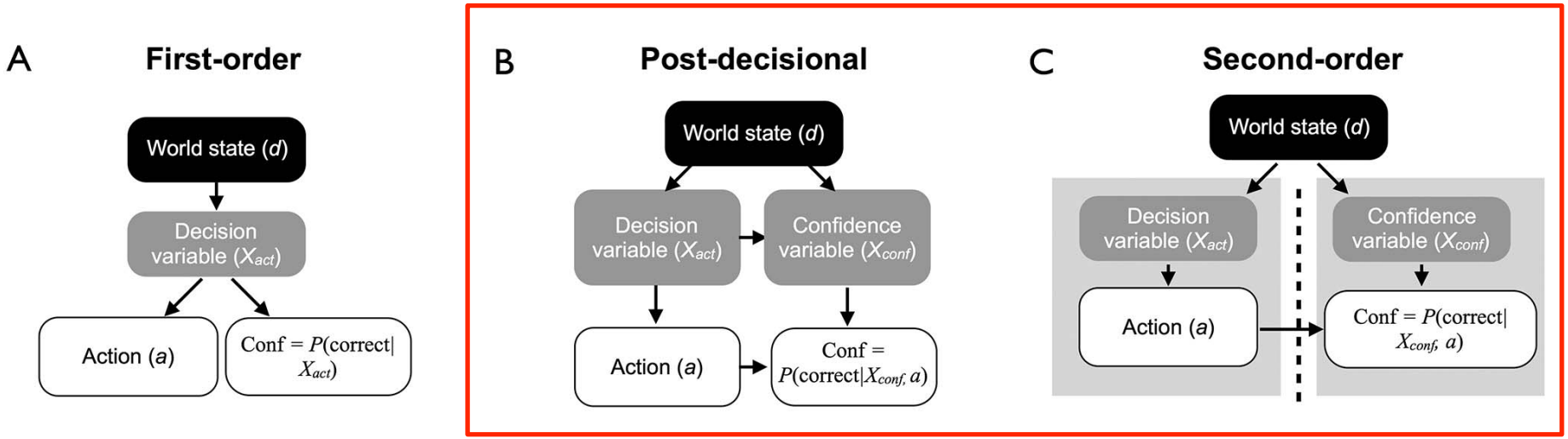
?

Confidence = *P(d=a|data)*

L or R

Action (*a*)

Computational building blocks of metacognition

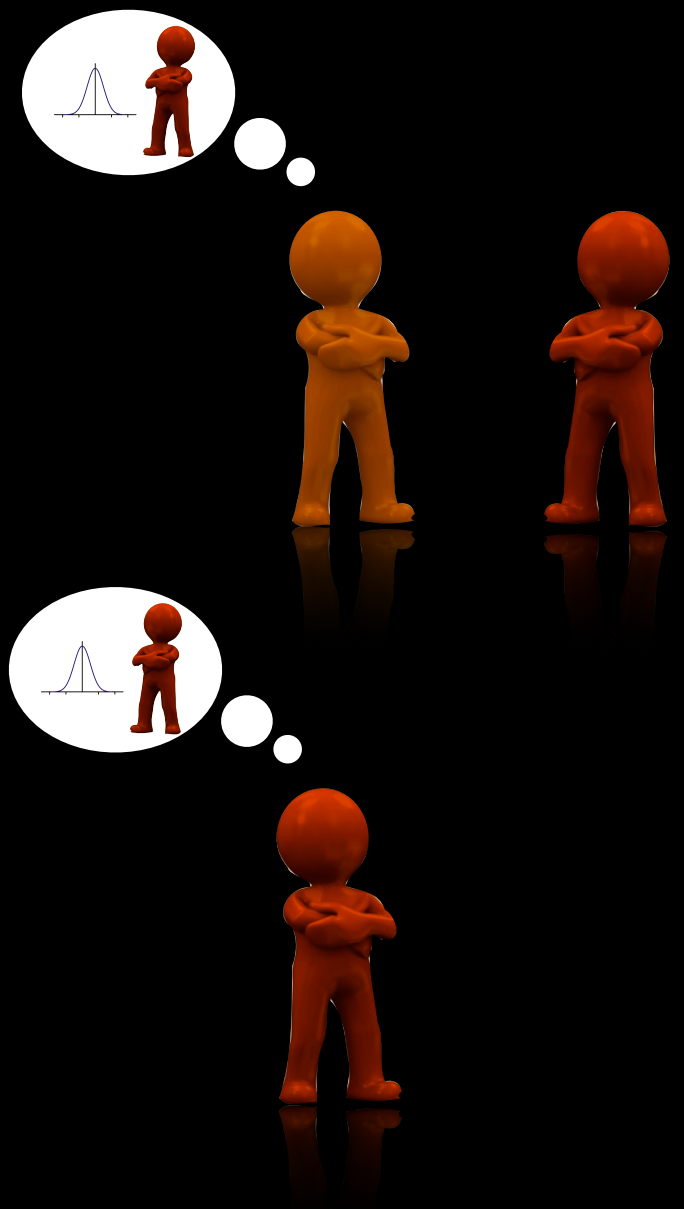


Second-order models a) permit dissociation between performance and metacognition and b) predict confidence is a late-stage construction from state/action variables

Fleming & Daw (2017) *Psych Rev*

Second-order computation

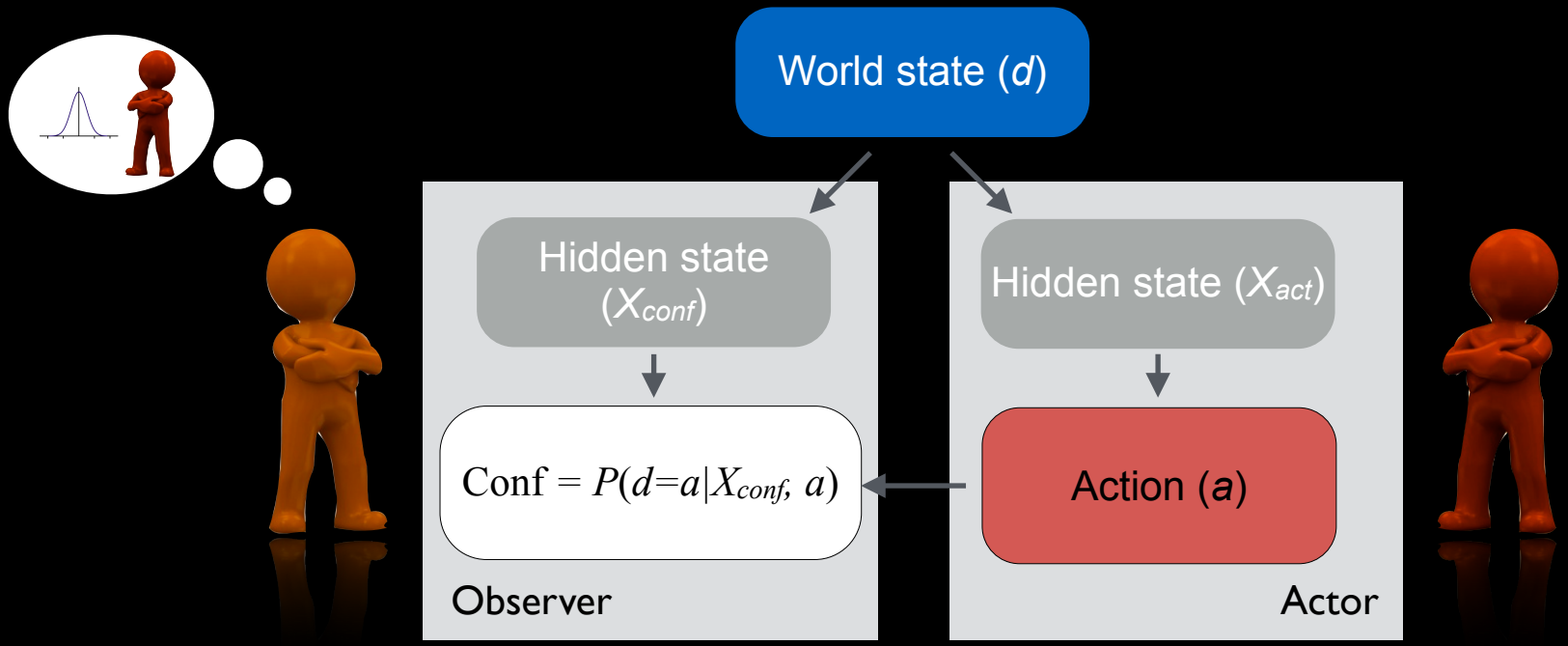
Estimating one's own confidence is computationally equivalent to estimating the performance of another individual



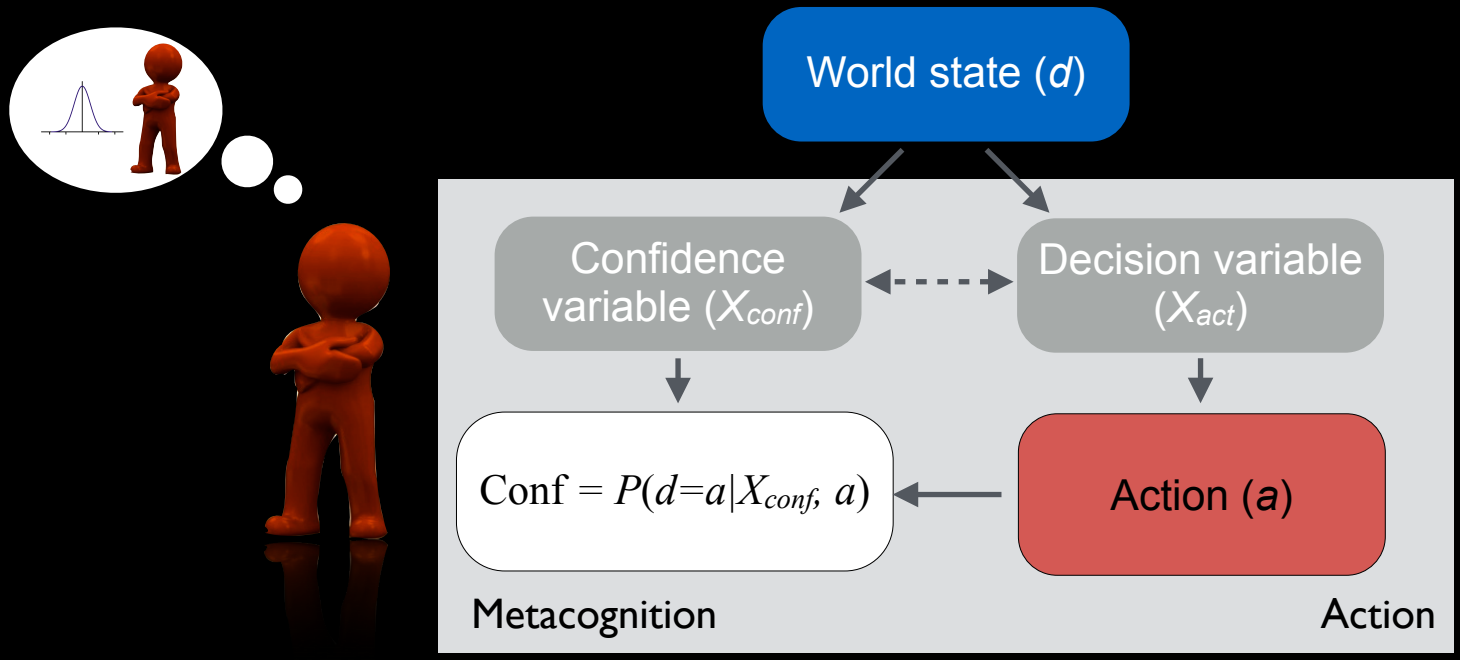
1) Between-subject case

2) Within-subject case

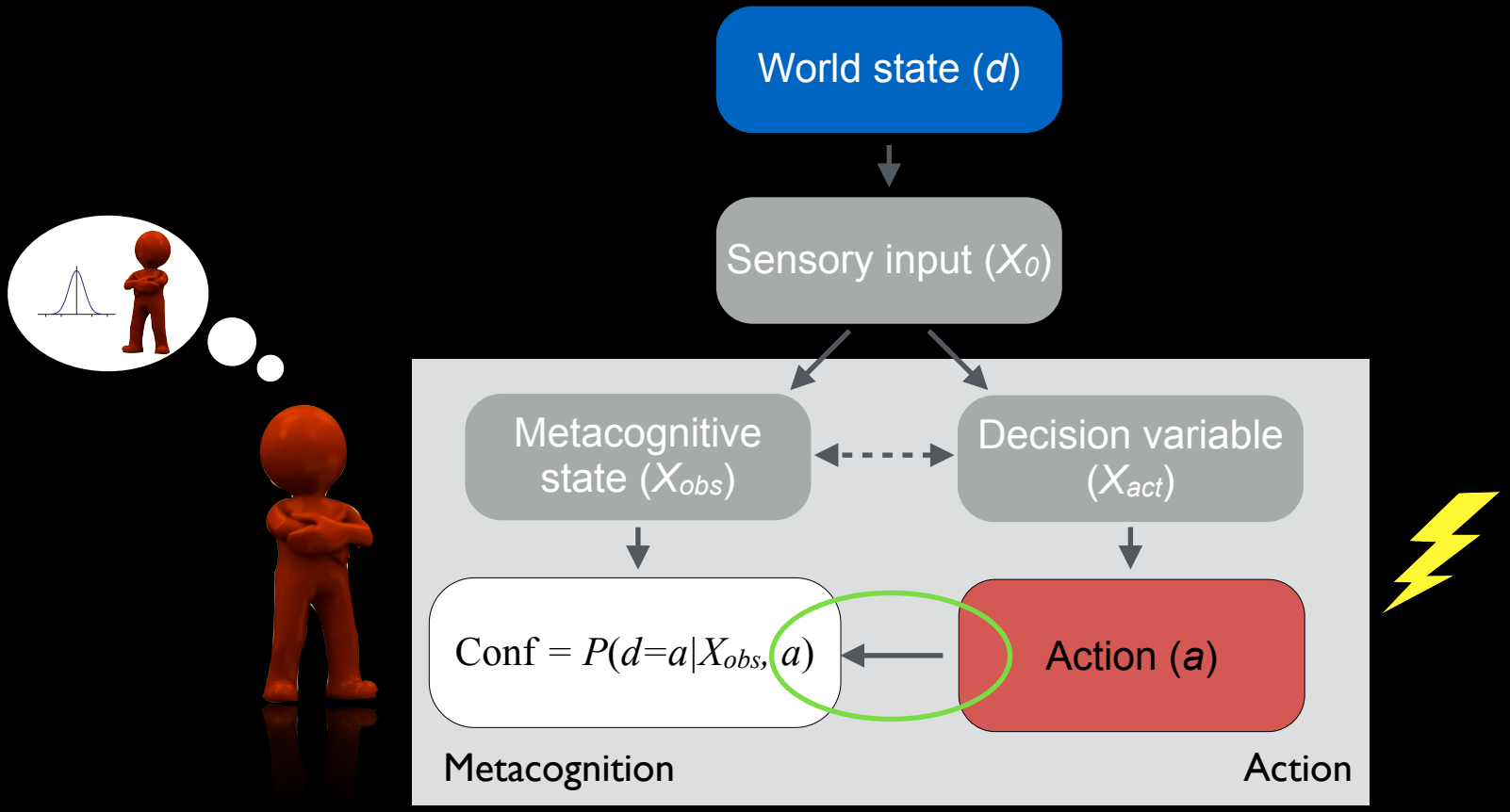
Second-order computation - between-subject case



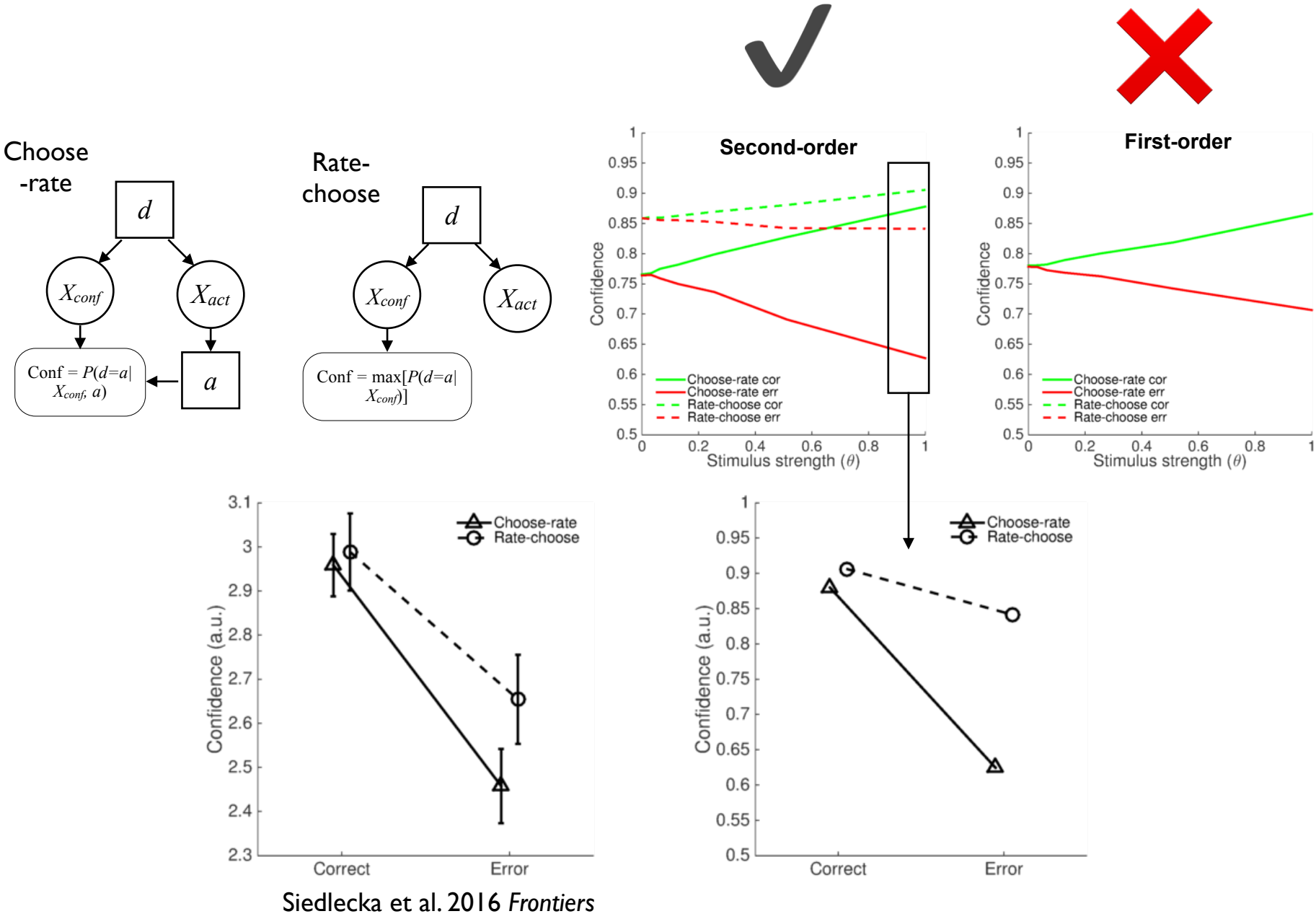
Second-order computation - within-subject case



Influence of motor system on confidence reports

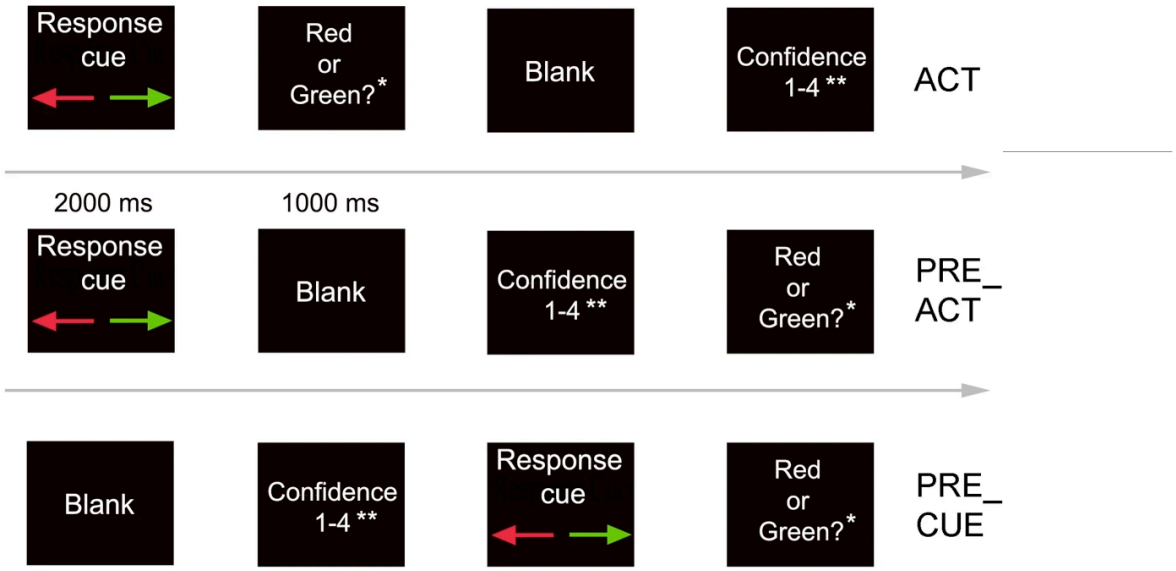


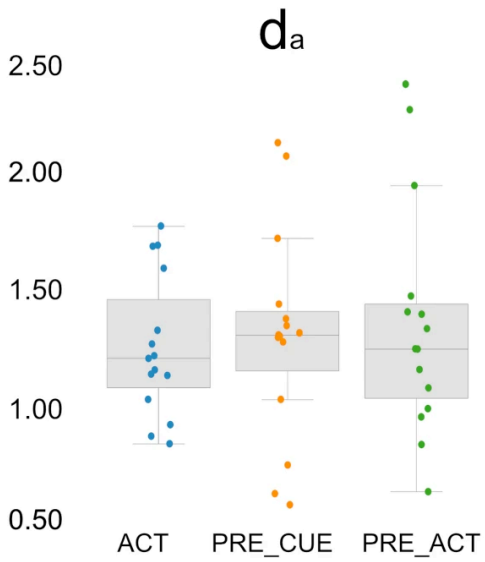
Paradoxical effects of self-action on metacognition

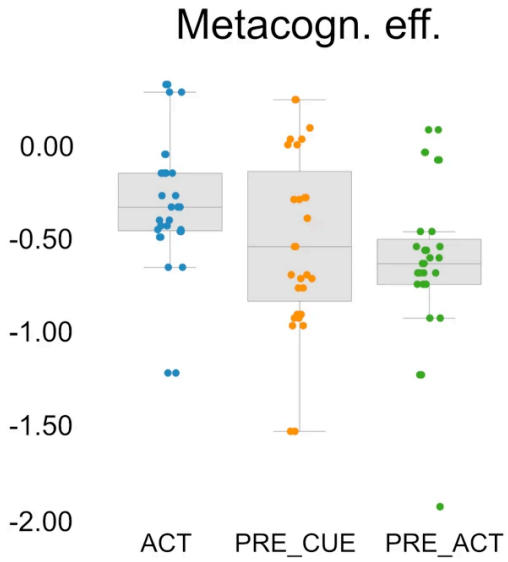


Siedlecka et al. 2016 *Frontiers*

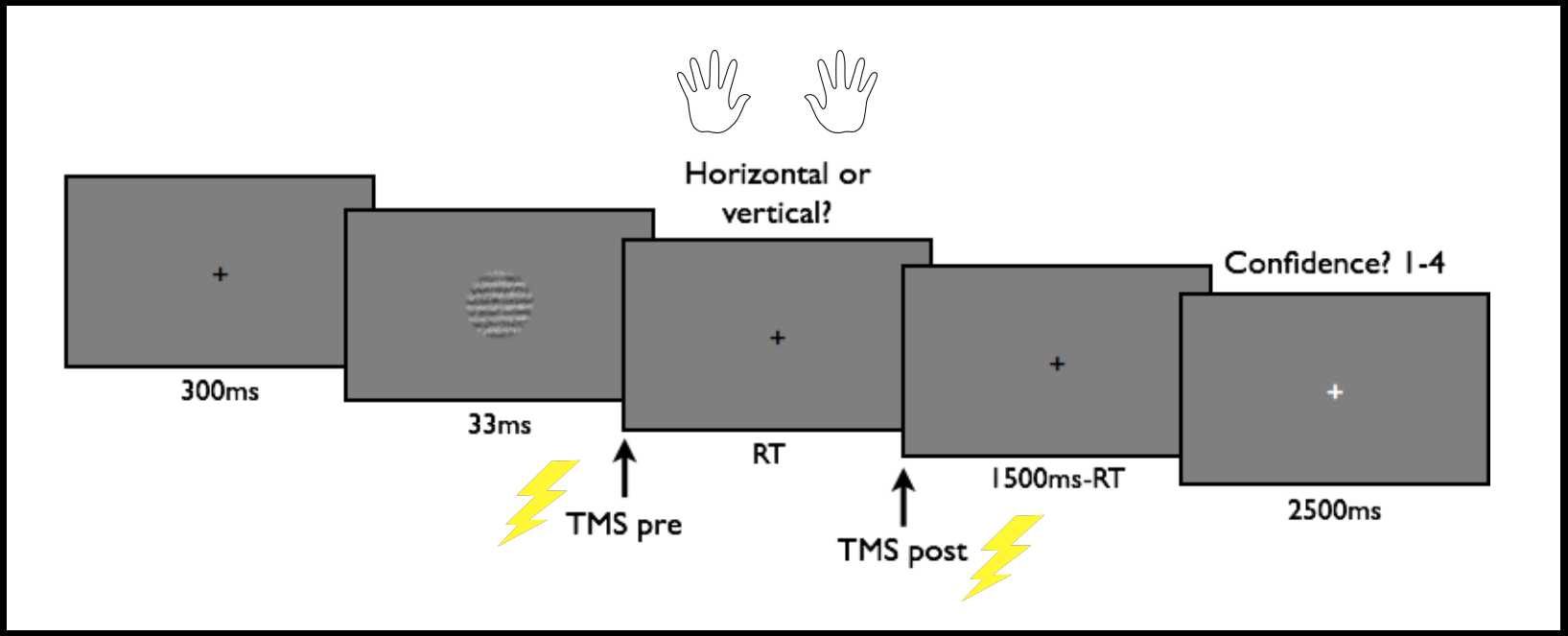
Paradoxical effects of self-action on metacognition

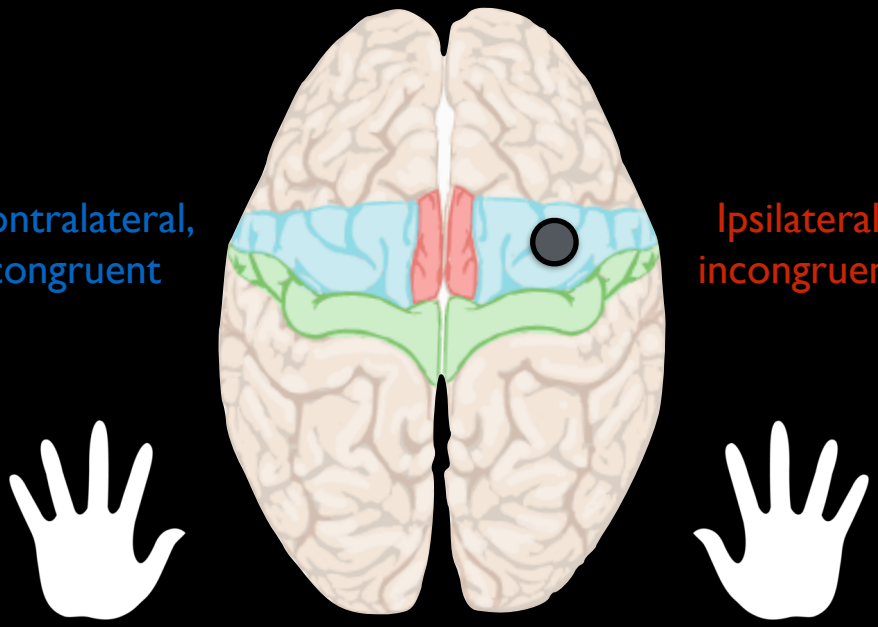






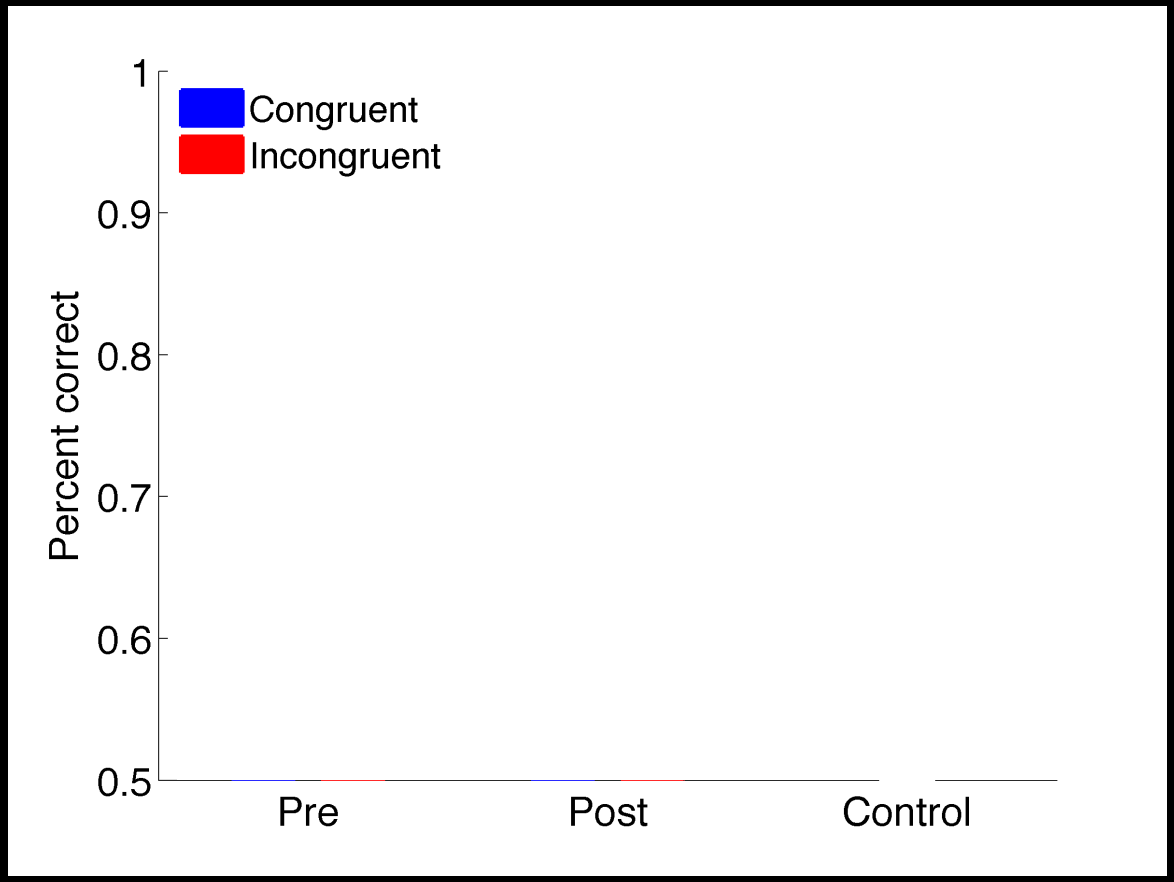
Wokke et al. 2020 *Sci Reports*





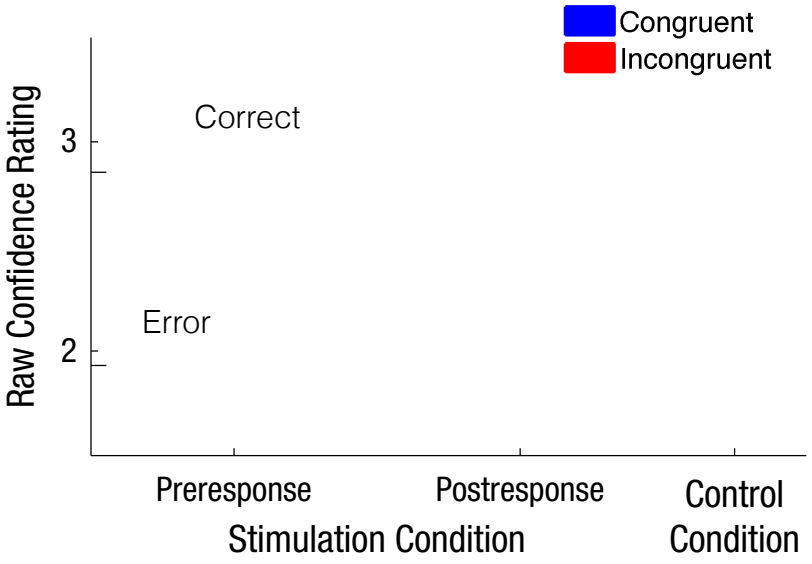
Fleming, Maniscalco, Ko, Amendi, Ro & Lau (2015) *Psych Science*

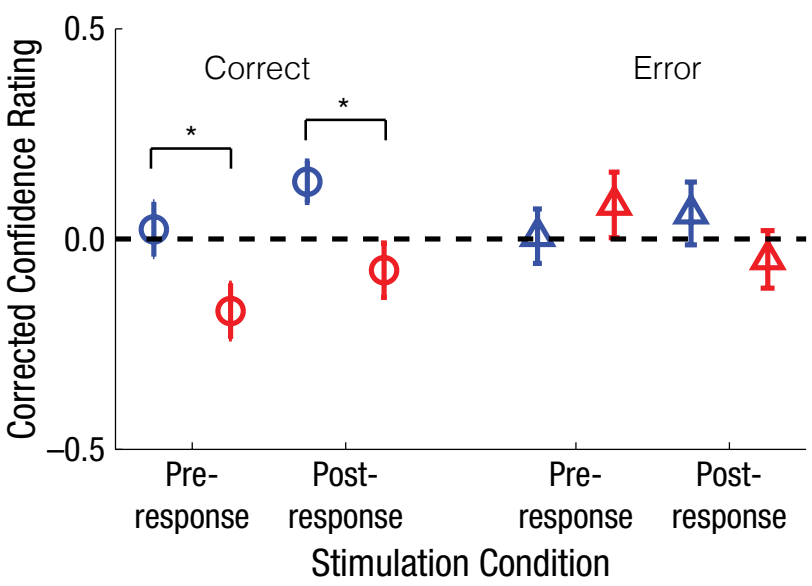
Influence of motor system on confidence reports



Fleming, Maniscalco, Ko, Amendi, Ro & Lau (2015) *Psych Science*

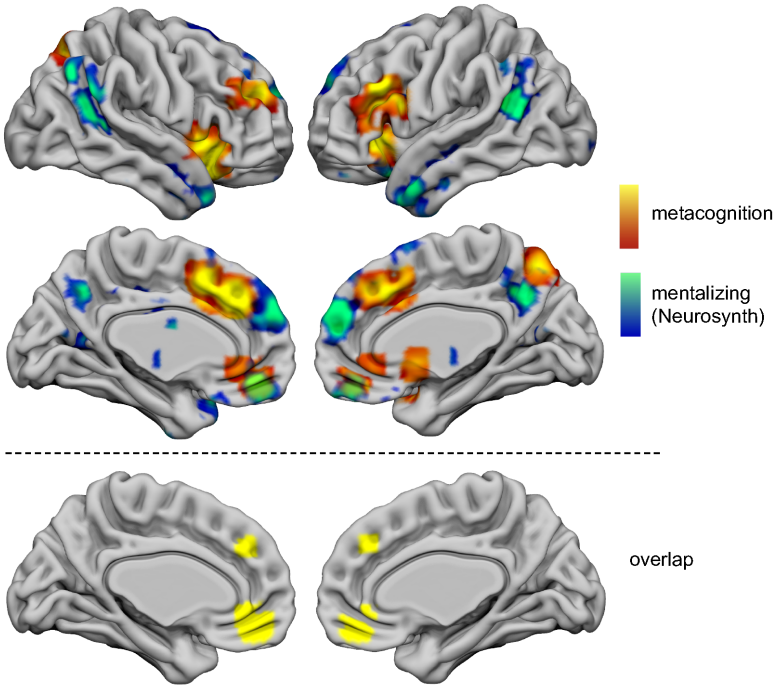
Influence of motor system on confidence reports



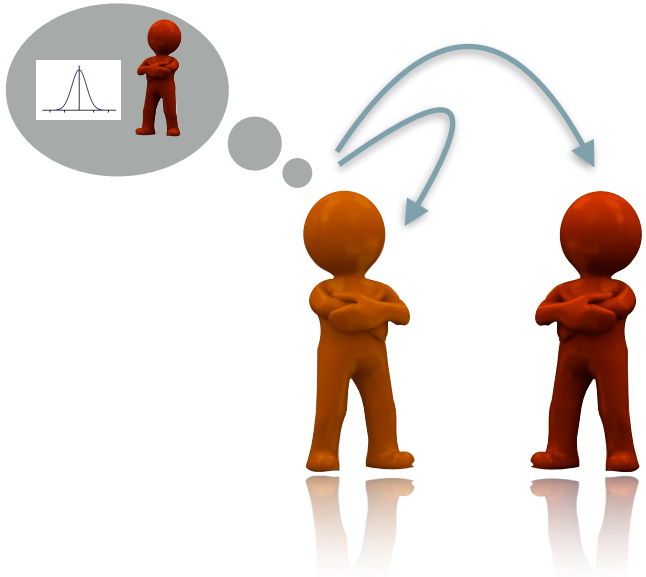


Fleming, Maniscalco, Ko, Amendi, Ro & Lau (2015) *Psych Science*

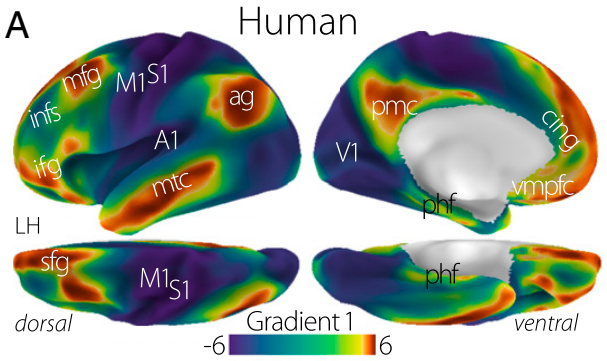
Brain networks for thinking about thinking

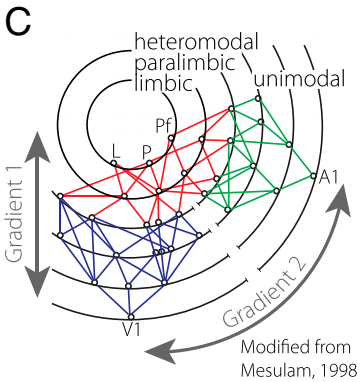


Vaccarro & Fleming (2018) *Brain Neuroscience Advances*

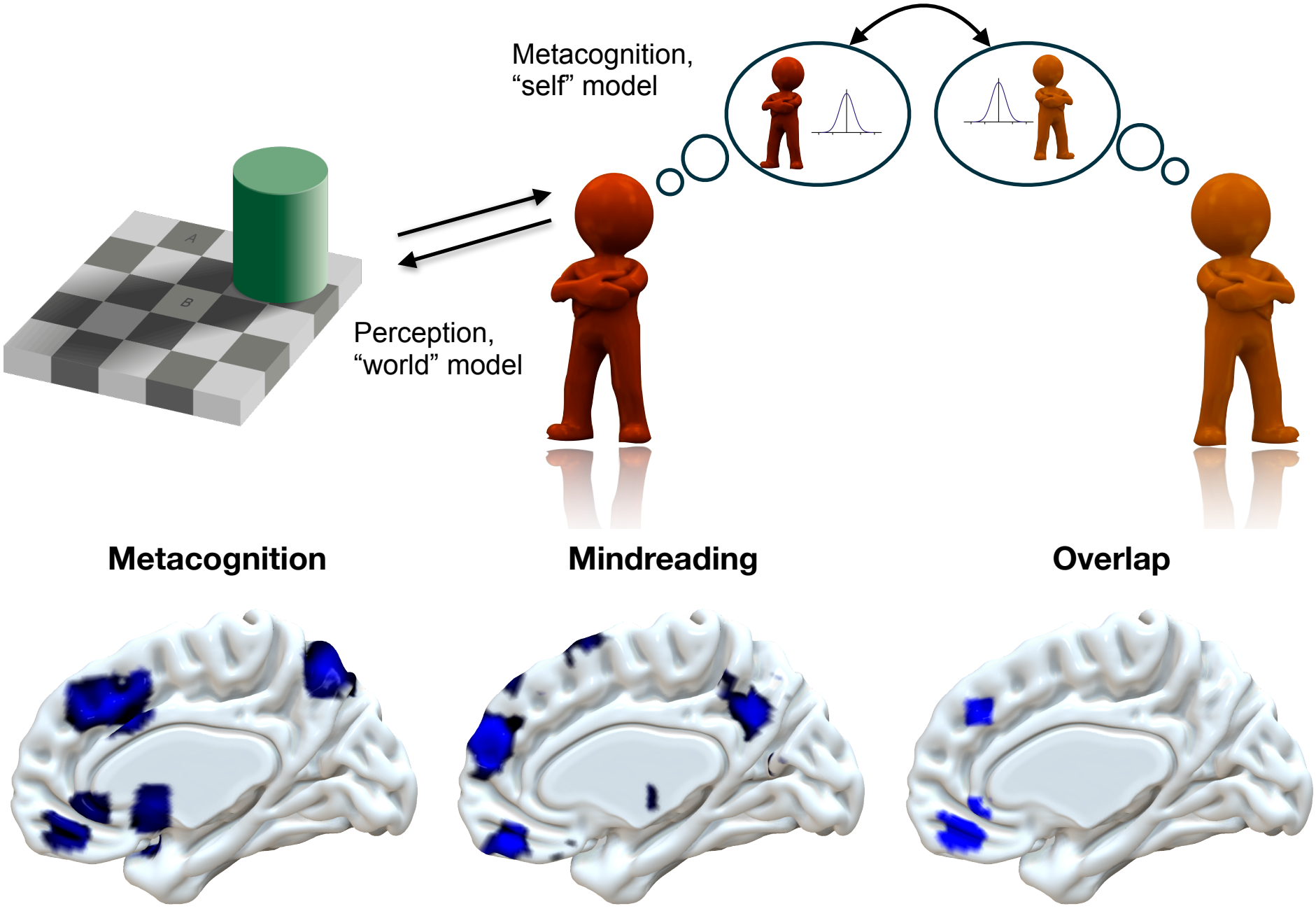


Fleming & Daw (2017) *Psych Rev*





Margulies et al. (2016) *PNAS*


Metacognition,
“self” model
Perception,
“world” model
A
B
Metacognition
Mindreading
Overlap

Linking Metacognition and Mindreading: Evidence From Autism and Dual-Task Investigations

Toby Nicholson and David M. Williams
University of Kent

Sophie E. Lind
City, University of London

Catherine Grainger
University of Stirling

Peter Carruthers
University of Maryland

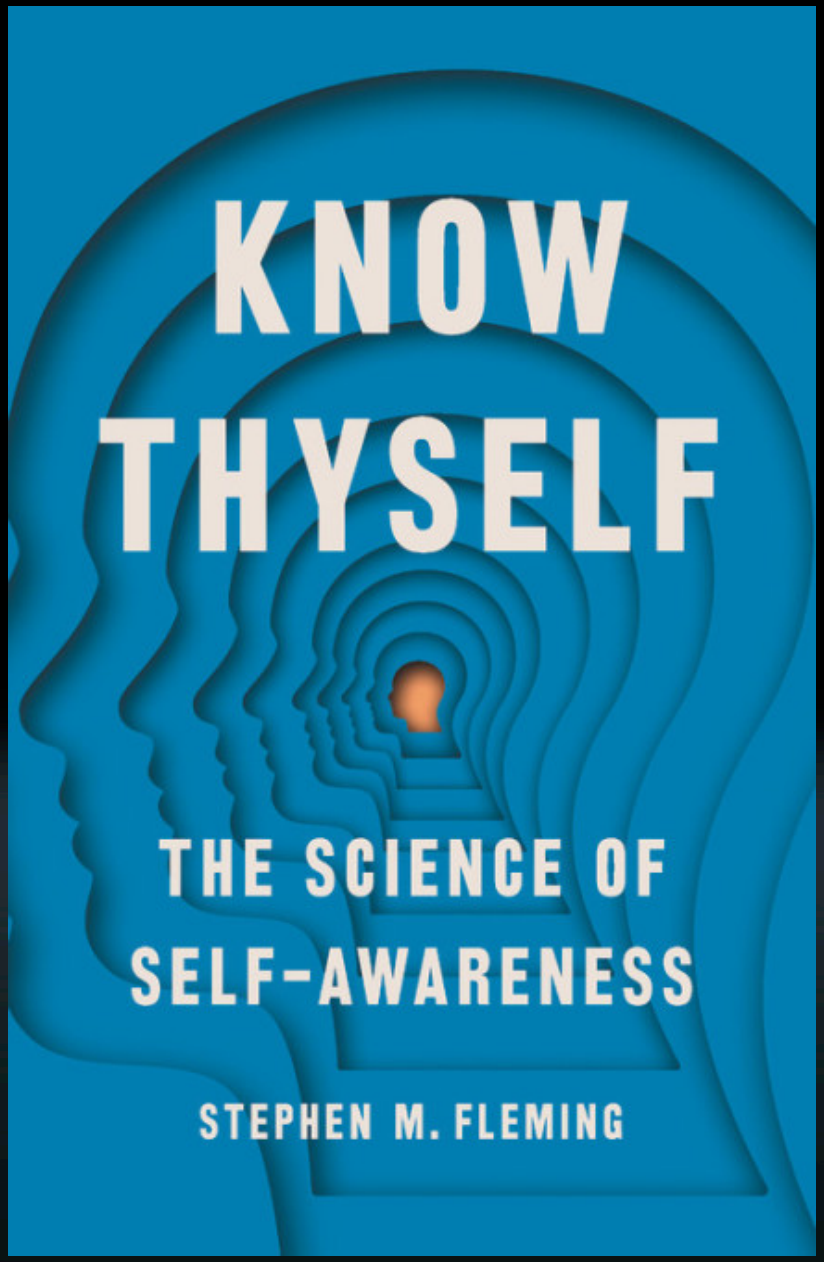
- Autistic participants were impaired in explicit confidence ratings (but not implicit gambles) compared to neurotypical participants
- No differences in first-order task performance between ASD and NTs
- A secondary task that involves thinking about others ("theory of mind" task) interferes with explicit (but not implicit) metacognition about self
- A similarly demanding secondary task that does *not* involve thinking about others does not interfere with metacognition

Nicholson et al. (2020) *JEP:General*

"How should I know what I'm thinking? I'm not a mind reader."

Summary

- We can measure metacognition across different tasks as the statistical association between behaviour and self-evaluation (confidence)
- Adopting a signal detection theory framework allows simultaneous estimation of both first-order (d') and metacognitive (meta-d') sensitivity
- Metacognitive confidence is encoded in activation patterns in PFC independently of behavioural performance
- Human-level self-awareness may be supported by second-order computations that share resources with the capacity to think about others (theory of mind)



metacoglab.org/book

@smfleming