Measuring consciousness: From behaviour to neurophysiology

Anil Seth
University of Sussex, Brighton, UK
www.anilseth.com

Liege, April 2009

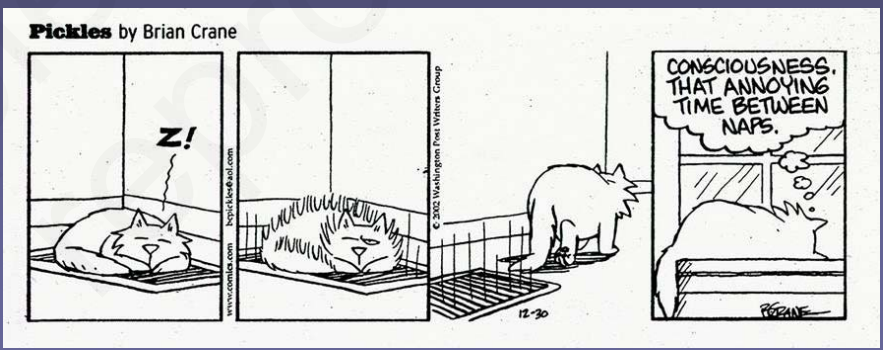
Outline

- Consciousness
- Behavioral measures
  - Example: Post-decision wagering
- Brain-based measures
  - Example: Complexity and causal density
- Combining multiple measures, conflicts and synergies
- (boundaries of consciousness)

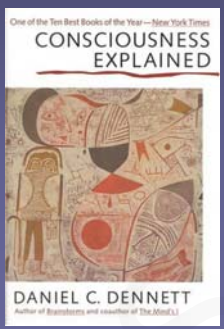
Consciousness

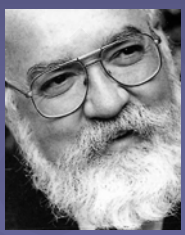
"Consciousness is everything we experience. Think of it as what abandons us every night when we fall into a dreamless sleep and returns the next morning when we wake up."

Tononi & Edelman (1998)



"Nobody has the slightest idea how anything material could be conscious. *Nobody even knows what it would be like to have the slightest idea how anything material could be conscious.*"





Jerry Fodor (1992)

Consciousness

- Primary consciousness: basic components of conscious scenes: colors, shapes, smells, sounds.
- Higher-order consciousness: consciousness *of* consciousness: thoughts, beliefs, etc.

"tomato!"

Measuring consciousness

- Having a dependable measure(s) of consciousness is vital for a mature science of consciousness.
- Certain measures presuppose certain theories, and certain theories recommend the use of particular measures.

Theories of consciousness

- 'Wordly discrimination theories' (WDT)
  - Consciousness expressed in ability to discriminate.
- Integration theories (IT)
  - Consciousness reflects integration of otherwise independent cognitive and neural processes.
- Higher-order-thought (HOT) theories:
  - A mental state is conscious in virtue of the existence of a "higher-order" thought, distinct from that state, to the effect that one is in that state.

Seth et al (2008), *Trends Cog Sci*

Behavioural measures

Behavioural measures

- Objective measures: the ability to choose accurately under forced choice conditions.
- Strategic control: the ability to use or not use knowledge according to instructions (e.g., Jacoby).
- Subjective measures: ascertain whether subjects know that they know (introspection, confidence ratings, etc.).
- Recent measures: e.g., post-decision wagering (Persaud et al.,2007).

Seth (2008), *Consc. Cogn.*

Seth et al (2008), *Trends Cog Sci*

Behavioural measures

- Objective measures: the ability to choose accurately under forced choice conditions. WDT
- Strategic control: the ability to use or not use knowledge according to instructions (e.g., Jacoby). IT
- Subjective measures: ascertain whether subjects know that they know (introspection, confidence ratings, etc.). HOT
- Recent measures: e.g., post-decision wagering (Persaud et al.,2007). HOT

Seth (2008), *Consc. Cogn.*

Seth et al (2008), *Trends Cog Sci*

Post-decision wagering

Post-decision wagering (PDW)

- "A new objective measure of awareness" [which avoids] "the uncertainties associated with the conventional subjective measures of awareness (verbal reports and confidence ratings)"
- PDW "measures awareness directly"

Persaud et al (2007), *Nat Neuro*

Post-decision wagering (PDW)

- Subjects make a 'first order' discrimination.
- They then place a (high or low) wager on the correctness of this discrimination.
- If they believe they are guessing, they should wager low (or be indifferent).
- If they have any confidence, they should wager high.
- Examples: Blindsight in GW, Iowa gambling task
- History: Ruffman et al. (2001); Shields et al. (2005)

Shields et al (2005), *J. Gen. Psychol.*

Ruffman et al (2001), *J. Exp. Chi. Psychol.*

Theoretical objections

- Absence of evidence is not evidence of absence (unless you accept HOT).
- PDW is if anything *more* indirect than confidence ratings: is it possible to learn implicitly to wager advantageously?
- All behavioral measures have a response criterion potentially subject to bias. For PDW it is risk aversion.
- PDW highlights the interdependence of measures and theories.

Seth (2008a,b), *Consc Cogn*

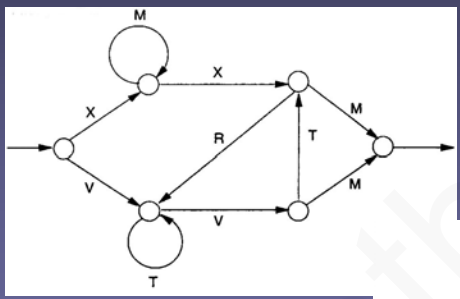
Seth et al (2008), *Trends Cog Sci.*

PDW and confidence ratings (CR)

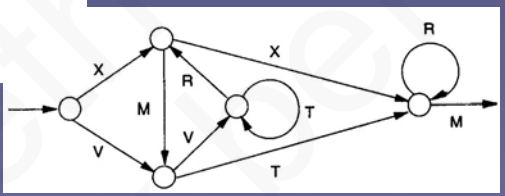
- CR bias: subjects may *think* they know to some degree but *say* they know nothing at all.
- PDW bias: subjects may *think* they know to some degree but still *wager low* in order to avoid losses (loss/risk aversion)
- In practice, which is more sensitive?

PDW in an artificial grammar paradigm

- Subjects are trained and tested on a standard AGL paradigm.



A



B

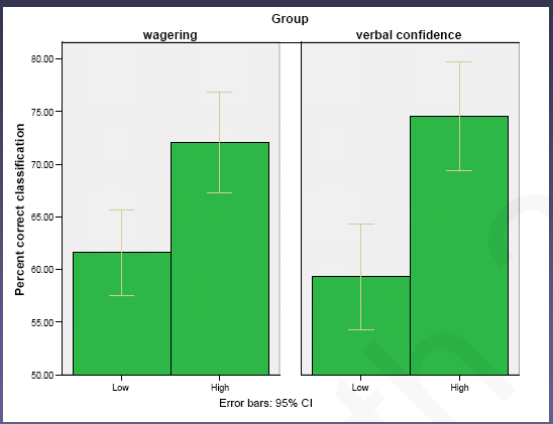
Dienes & Seth (in review), *Consc Cogn*

PDW in an artificial grammar paradigm

- 50% of subjects are asked to rate their choices via binary CR, and 50% via wagering (with sweets as reward).
- All subjects are given a risk-aversion questionnaire (Hartog et al., 2000).

Dienes & Seth (in review), *Consc Cogn*

PDW in an artificial grammar paradigm



No difference in sensitivity between CR and PDW

- The more risk averse a person was, the lower the measured amount of conscious knowledge used during PDW (but not during CR).

Dienes & Seth (in review), *Consc Cogn*

PDW in an artificial grammar paradigm

- PDW is not more sensitive than CR as a measure of consciousness in this paradigm.
- Subjects were more likely to indicate *some* confidence using CR than using PDW.
- PDW but not CR depends on individual differences in risk aversion.
- We also introduce a 'no-loss' version of PDW which eliminates risk-aversion biases.

Dienes & Seth (in review), *Consc Cogn*

Brain-based measures

Brain-based measures

- Low amplitude, irregular EEG activity during waking (Berger, 1929); bispectral index (BIS).
- ERPs (e.g., early or late visual evoked potentials).
- Widespread activation
- Synchrony (e.g., γ band, β band).
- Dynamical complexity measures

Seth et al (2008), *Trends Cog Sci*

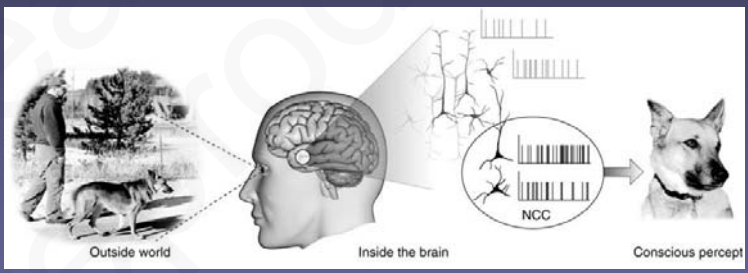
Brain-based measures

- Low amplitude, irregular EEG activity during waking (Berger, 1929); bispectral index (BIS). IT?
- ERPs (e.g., early or late visual evoked potentials). IT?
- Widespread activation IT
- Synchrony (e.g., γ band, β band). IT?
- Dynamical complexity measures. IT

Seth et al (2008), *Trends Cog Sci*

Correlates of consciousness

- Neural correlates: activity in groups of neurons or brain regions that has a privileged relationship with consciousness.



Koch (2007), *Scholarpedia*

- Explanatory correlates: brain processes that account for fundamental (structural) aspects of conscious experience.

Seth (2009). *Cognitive Computation*

Structural properties of consciousness

- Aspects or dimensions of the way the world is presented to us through conscious experience:
- Simultaneous integration and differentiation (dynamical complexity)
- Perspectivalness
- Emotion and mood
- Volition/intentionality

Seth (2009). *Cognitive Computation*
Seth & Clowes (2008). *Art. Intell. In Medic.*

Structural properties of consciousness

- Aspects or dimensions of the way the world is presented to us through conscious experience:
- Simultaneous integration and differentiation (dynamical complexity)
- Perspectivalness
- Emotion and mood
- Volition/intentionality

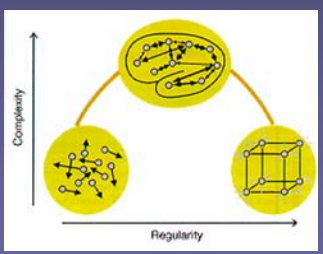
Seth (2009). *Cognitive Computation*
Seth & Clowes (2008). *Art. Intell. In Medic.*

Complexity and causal density

Dynamical complexity

Every conscious scene is *differentiated*

Every conscious scene is *integrated*

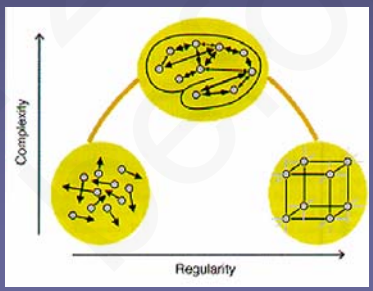


Tononi & Edelman (1998), *Science*

Edelman & Tononi (2000), *A Universe of Consciousness*

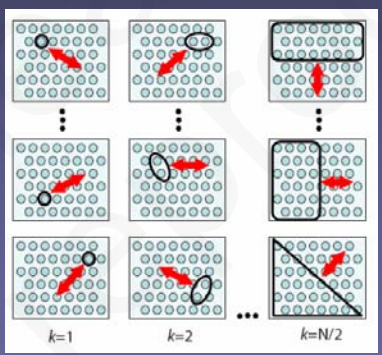
Measures of dynamical complexity

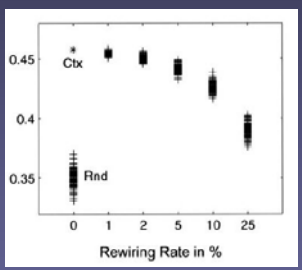
- Neural complexity (Tononi, Sporns, Edelman, 1994)
- Information integration (Φ) (Tononi, 2004)
- Causal density (Seth, 2005, 2008)



Neural complexity

- Small parts of a system are independent, large parts are comparatively integrated.





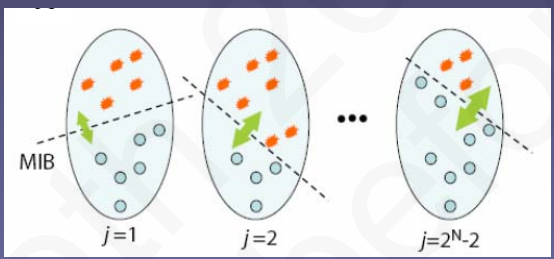
$$C_N(\mathrm{X}) = \sum_k \langle MI(\mathrm{X}_j^k; \mathrm{X} - \mathrm{X}_j^k) \rangle,$$

Tononi, Sporns, & Edelman (1994), *Proc. Nat Acad. Sci. USA*

Sporns, Tononi, & Edelman (2000), *Cereb. Cort.*

Information integration (Φ)

- 'Effective information' across the 'informational weakest link' (MIB).
- Φ is measure of the capacity of a system to integrate information.

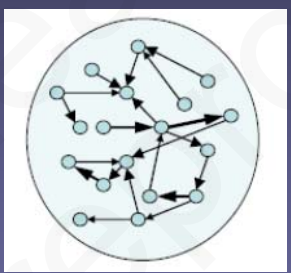


$$\Phi = min\,(\mathrm{EI}(A \rightarrow B) + \mathrm{EI}(B \rightarrow A))\,.$$

Tononi (2004), *BMC Neuroscience*

Causal density

- Total amount of causal interactivity in a system.



$$c_{dw}(\mathrm{X}) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} g c_{j \rightarrow i}$$

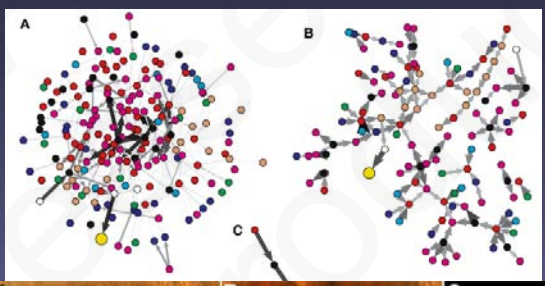
Seth (2005), *Network: Comp. Neur. Sys.*

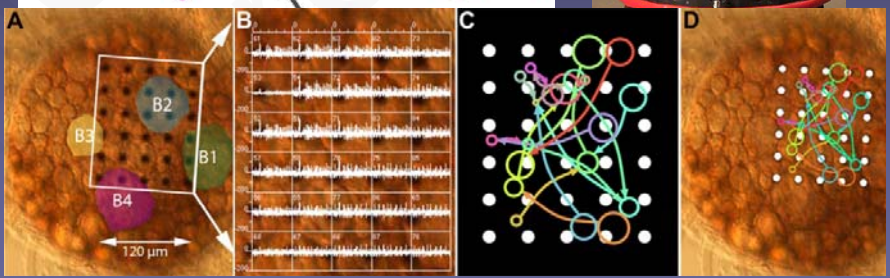
Seth (2008) *Cogn. Neurodyn.*

Granger (G-) causality

Granger (1969), *Econometrica.*

G-causality in practice



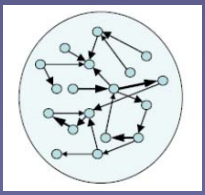


Passaro, Seth, et al. (in preparation)

Seth & Edelman. (2007). *Neur Comp.*

Causal density

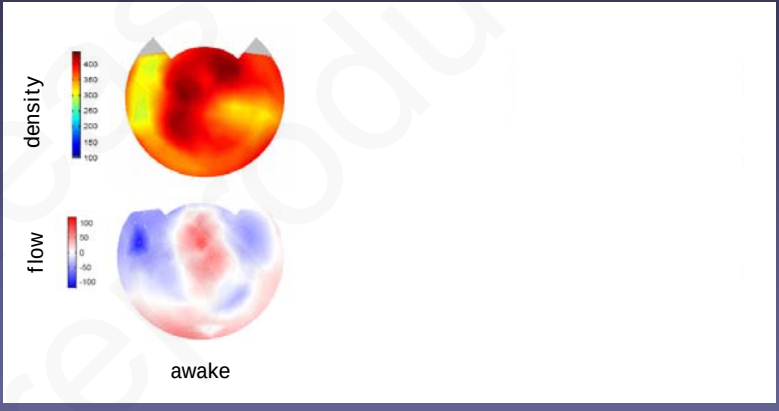
- Independent elements will have low causal density, as will elements that behave identically.
- Each subset *must behave differently* from others, in order to contribute new predictive information; each subset *must be integrated* with other subsets, in order for this information to be useful.



Seth (2005), *Network: Comp. Neur. Sys.*

Seth (2008) *Cogn. Neurodyn.*

Causal density in MEG data



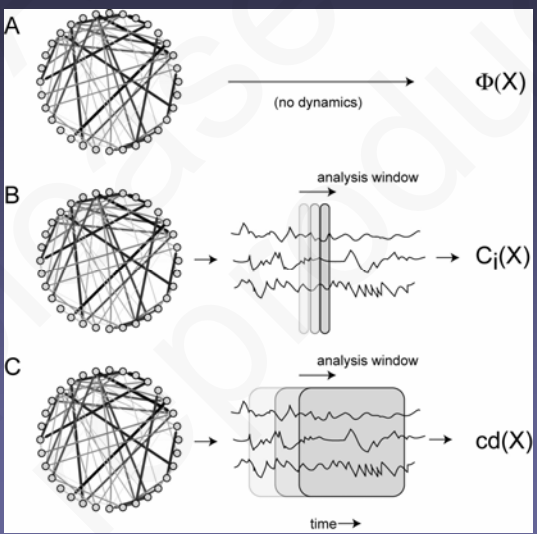
Seth (2007). Soc. Neuro. Abs.

Why different measures?

- Different measures can operationalize subtly different aspects of the same overarching property:
  - Unlike Φ, causal density and neural complexity are sensitive to the activity and not the capacity of a system.
- Different measures can correct perceived deficiencies:
  - Unlike neural complexity, Φ and causal density are sensitive to causal interactions.
  - Unlike Φ, causal density and (approximate) neural complexity can be measured for non-trivial systems.

Seth, A. et al (2006). *Proc. Nat. Acad. Sci.*

Incorporation of time



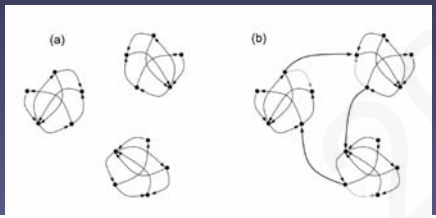
Unlike Φ and neural complexity, causal density is sensitive to neural dynamics that are 'smeared out' over time.

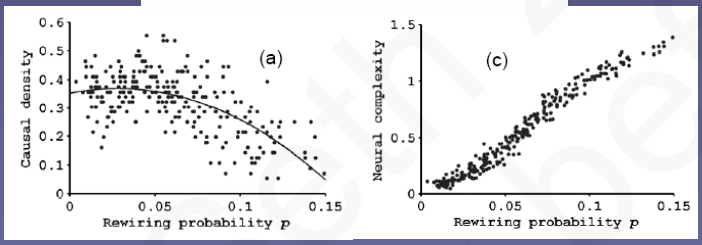
Seth, A (2009). *Cogn. Comput.*

Seth, A. et al (2006). *Proc. Nat. Acad. Sci. USA*

Causal density in models

Causal density behaves better than neural complexity when tested on small-world networks of spiking neurons.





Shanahan, M.P. (2008). *Phys. Rev. E.*

Conflicts and synergies between measures

Conflicts between measures

Content can be conscious according to 'widespread activation' but unconscious according to subjective measures.

Unconscious Activation of the Cognitive Control System in the Human Prefrontal Cortex

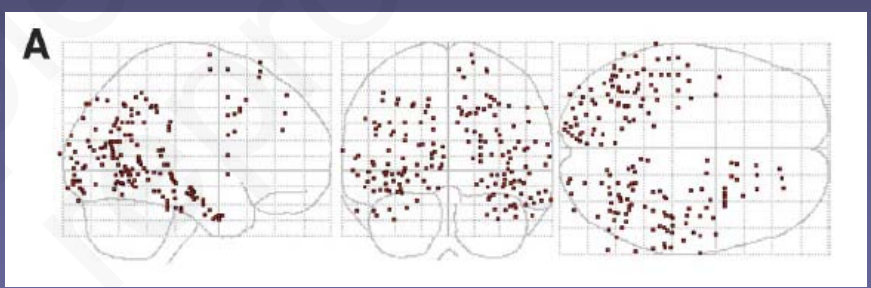
Hakwan C. Lau[1,2] and Richard E. Passingham[1,2]
[1]Wellcome Department of Imaging Neuroscience, University College London, London WC1N 3BG, United Kingdom, and [2]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom

Lau & Passingham (2007), *J. Neurosci*

Integrating multiple measures

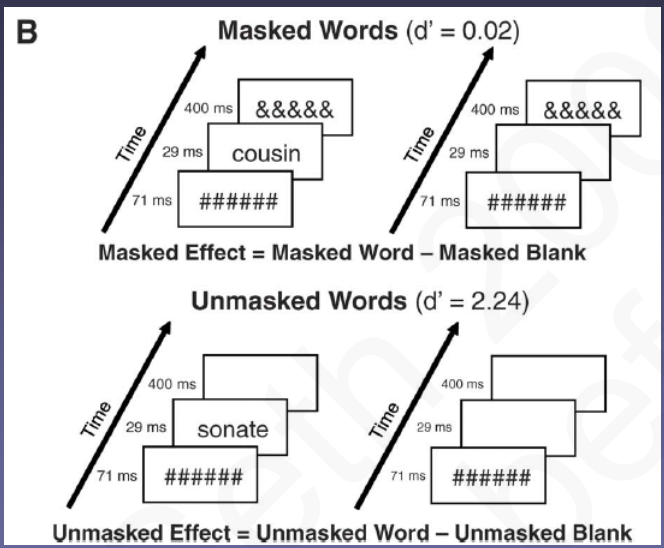
Converging Intracranial Markers of Conscious Access

Raphaël Gaillard[1,2,3], Stanislas Dehaene[1,4,5], Claude Adam[6], Stéphane Clémenceau[6], Dominique Hasboun[6,7], Michel Baulac[6,7], Laurent Cohen[1,6,7], Lionel Naccache[1,6,7*]



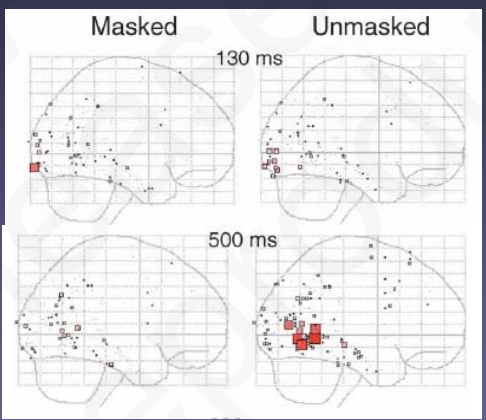
Gaillard et al. (2009), *PLOS Biol.*

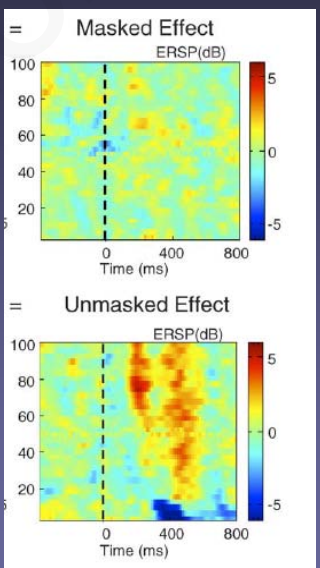
Integrating multiple measures



Gaillard et al. (2009), *PLOS Biol.*

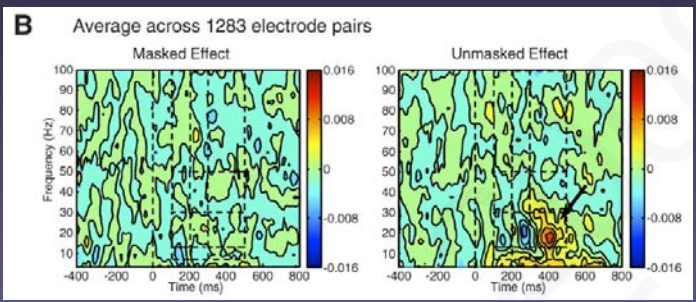
Integrating multiple measures

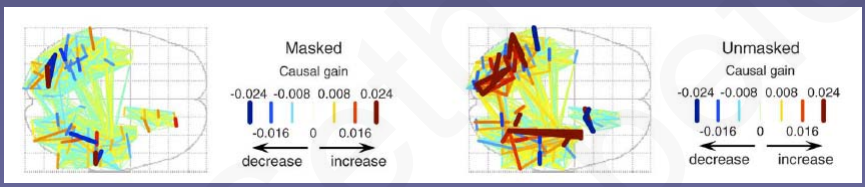




Gaillard et al. (2009), *PLOS Biol.*

Integrating multiple measures





Gaillard et al. (2009), *PLOS Biol.*

Integrating multiple measures

- Behavioral measures: forced-choice discrimination, verbal report.
- Brain measures: ERPs, event-related spectral perturbations (ESRP), phase coherence, Granger causality
- Conflicts:
  - increase in beta phase synchrony at same time as decrease in beta ESRP
- "conscious processing … can be reflected by many partially overlapping physiological measures"

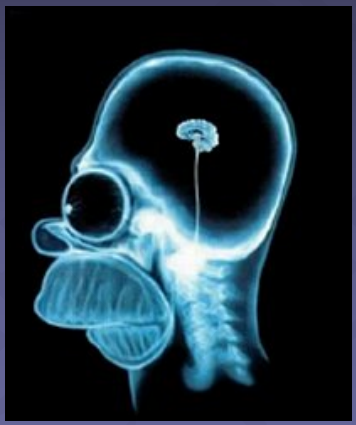
Gaillard et al. (2009), *PLOS Biol.*

Pitfalls in designing and applying measures

Pitfalls

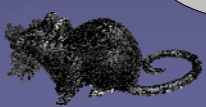
- A mismatch between application and intuition could require us to:
- Update our intuition about what should be measured (i.e., what is ‘complexity’ anyway)?
- Update our implementation of the measure; does the measure capture adequately our intuition?
- Update our methodology by which the measure is applied to data; e.g., is the code correct, are assumptions on the data satisfied?

The boundaries of consciousness



Animal consciousness

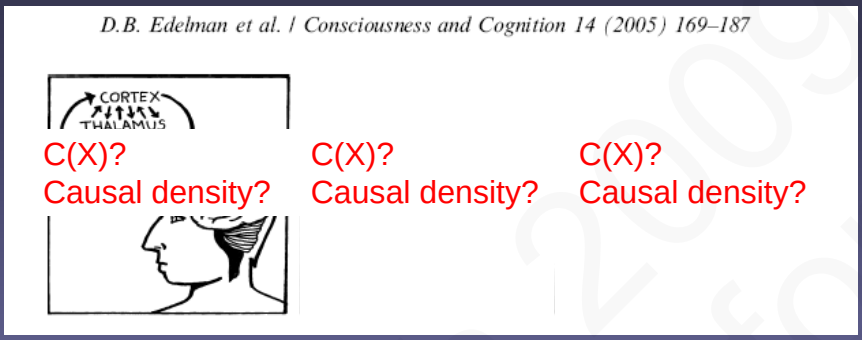
"tomato!"



"_____"

Seth, Baars, & Edelman D.B. (2005). *Consc. Cogn.*

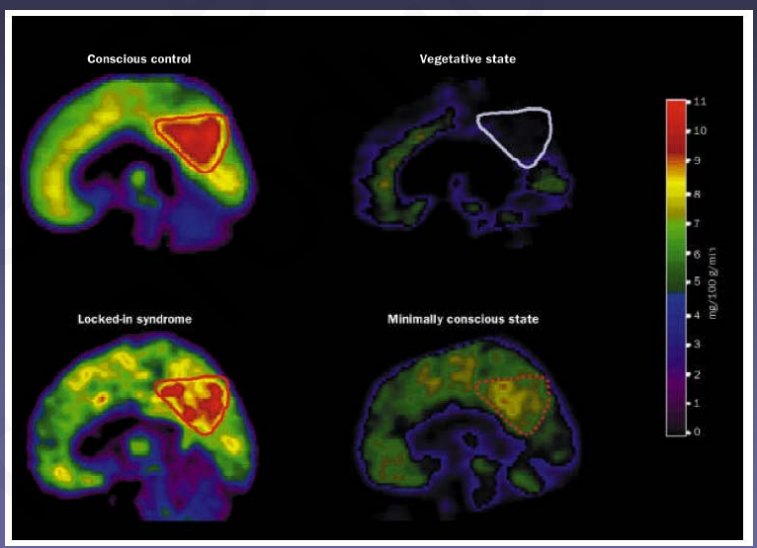
Animal consciousness



Edelman, D.B., Baars, & Seth (2005). *Consc. Cogn.*

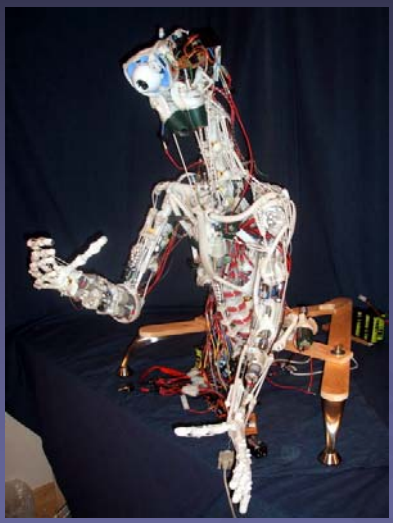
Seth, Baars, & Edelman D.B. (2005). *Consc. Cogn.*

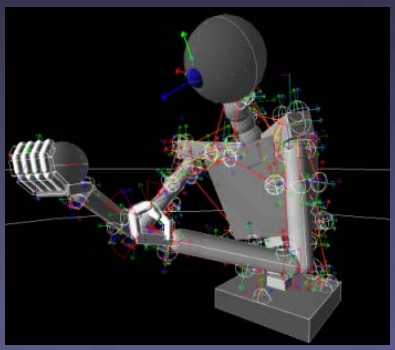
Residual consciousness



Laureys et al. (2004). *The Lancet*

Machine consciousness





Holland (2007). *J. Consc. Stud.*

Seth (2009). *Int. J. Mach. Consc.*

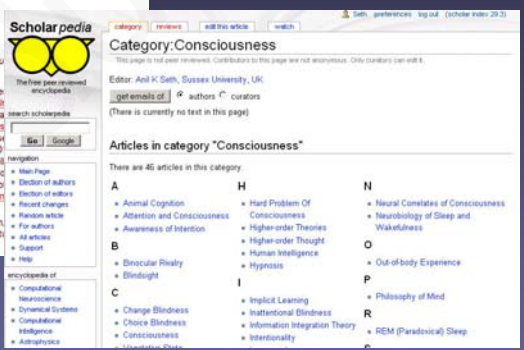
Summary (1)

- Behavioral measures: Hard to distinguish consciousness *per se* from reports of consciousness.
- Brain measures: Hard to ensure a measure has anything (much) to do with consciousness.
- Brain measures should operationalize explanatory correlates.
- Exciting new studies should combine multiple measures, behavioral and neurophysiological.

Summary (2)

- The ultimate virtue in a measure is not it's *a priori* robustness, but its ability to build on intuitions, identify interesting divides in nature, and correct the foundations on which it was built.

www.anilseth.com



www.scholarpedia.org/
Category:Consciousness

Sackler Centre for Consciousness Science