How can we Scientifically Study Consciousness?

David Gamez
Sackler Centre for Consciousness Science / Department of Informatics, University of Sussex, UK.
Turing Centenary Research Project, 3rd Workshop, Bucharest, 28th June 2015

Current State of Consciousness Research

- Lot of experimental work on the correlates of consciousness.
- Some mathematical theories.
- But discussions about consciousness regularly return to philosophical problems (zombies, colour inversion, causation, hard problem etc.).
- These are legitimate concerns that could undermine scientific results.

Objectives

- An approach to consciousness that is
  - Mathematical.
  - General.
  - Free from philosophical problems.

My Approach

- Framework of definitions and assumptions that handles or sets aside most philosophical problems with consciousness.
- Explains how we can accurately measure consciousness and the physical world.
- Enables us to develop and test mathematical theories of consciousness.
- Mathematical theories of consciousness can answer questions about the consciousness of bats, brain-damaged patients and robots.

Talk Overview

- Definition of consciousness.
- Science of consciousness:
  - Measurement of consciousness.
  - Measurement of the physical world.
  - Theories of consciousness.
  - Computational discovery of theories of consciousness.
- Applications
- Conclusion

Definition of Consciousness
**Bubbles of Experience**

- In our everyday encounters with the world we are immersed in a bubble of space centred on our bodies.
- I call this a bubble of experience.
- Bubbles of experience contain colours, smells, sounds, etc.

**Objective View**

**Bubble of Experience**

**Bubbles of Experience and the Physical World**

- Over the last 400 years there has been a gradual change in our interpretation of our bubbles of experience.
- Our theories about our bubbles of experience have evolved in response to our theories about the physical world.

**Naive Realism**

- We are naive realists in our day-to-day lives.
- Colours, smells, sounds etc. are properties of objects.
- Objects continue to have these properties when they are not being perceived.
**Atomism**
- Atomism was developed by the Ancient Greeks and revived in the 17th Century.
- Physical world consists of atoms and the void.
- Leads to a distinction between primary and secondary qualities:
  - **Primary qualities:**
    - Size, shape and motion.
    - Properties of atoms.
  - **Secondary qualities:**
    - Colours, smells, sounds, etc.
    - Appear when atoms interact with the senses.

**Bubbles of Experience and the Brain**
- Today the brain is interposed between our bubbles of experience and the physical world:
  - The physical world interacts with the senses.
  - Spiking patterns appear in the brain.
  - These spiking patterns are somehow linked to bubbles of experience containing colours, smells, etc.
- No longer any reason to believe that our bubbles of experience resemble the physical world.

**Modern Physics**
- In modern physics the physical world consists of wave-particles, superstrings, etc.
- We cannot imagine what these are like.
- The physical world has become a black box that is a source of signals (Russell).
- It no longer make sense to ask what the physical world is like.


Modern Physics

What is Consciousness?

- “Consciousness” is another name for bubbles of experience, which were reinterpreted in relation to an invisible physical world.
- Fits in with Wilkes’ claim that our modern use of the word ‘consciousness’ first appeared in the 17th Century.

Emergence of the Concept of Consciousness

Measurement of Consciousness

- Consciousness is measured through first-person reports.
- This raises a number of philosophical problems.
- These can be handled with assumptions.
- The science of consciousness is considered to be true given these assumptions.

Platinum Standard Systems

- To scientifically study consciousness we need a physical system that is known to be associated with consciousness.
- Cannot prove that a system is conscious.
- Platinum standard systems are systems that are assumed to be conscious.
- I have assumed that the normally functioning adult human brain is a platinum standard system.
CC sets

- A CC set is a set of spatiotemporal structures in the physical world that is correlated with a conscious state.
- This set is present when a conscious state is present and absent when the conscious state is absent.

CC sets are Functionally Connected to Consciousness

- A correlation between A and B is the same as a functional connection between A and B – they are different ways of stating that A and B deviate from statistical independence.
- There is a functional connection between a conscious state and its corresponding CC set.

Assumptions

1. During an experiment on consciousness, the consciousness associated with a platinum standard system is functionally connected to its reports.
2. During an experiment on consciousness all conscious states associated with a platinum standard system are available for report and all aspects of these states can potentially be c-reported.

Assumptions

3. The conscious state that is associated with a CC set nomologically supervenes on the CC set. In our current universe physically identical CC sets are associated with indistinguishable conscious states.
4. The normally functioning adult human brain is a platinum standard system.
5. The physical world is causally closed
6. CC sets cause a platinum standard system’s first person reports about consciousness.

Assumptions for the Measurement of Consciousness
Describing Conscious States

- Consciousness cannot be described in natural language:
  - Context-bound
  - Ambiguous
  - Not applicable to infants, bats, robots, etc.
  - Not mathematically tractable.

C-description

- Need a precise formal way of describing consciousness that is applicable to any system.
- This will be referred to as a c-description.
- Possible methods include:
  - XML/LMNL
  - High dimensional qualia
  - Category theory

Measurement of the Physical World

- The measured object interacts with a calibrated object.
- Observe result and extract a number.

Description of the Physical World

- The number that is extracted through a measurement procedure is attributed to an object in the physical world.
- 3 metres was the height of an elephant.
- Objects are tightly defined in physics and chemistry.
- They are not tightly defined in biology.
**P-description**
- We want a science of consciousness that can make predictions about the consciousness of arbitrary systems (bats, robots, rocks etc.).
- A science of consciousness based on biological neurons will not be able to say anything about the consciousness of systems based on synthetic neurons.
- Need a precise formal description of the spatiotemporal structures that form CC sets.
- Will be referred to as a *p-description*.

**Theories of Consciousness**
- A *c-theory* is a mathematical description of the relationships between *c-descriptions* and *p-descriptions*.
- It can generate *c-descriptions* from *p-descriptions*.
- It can generate *p-descriptions* from *c-descriptions*.

**Example: Tononi’s Information Integration Theory of Consciousness (IIT)**
- IIT is the closest thing to a *c-theory* that we have so far.
- A mathematical algorithm links a description of the physical world to a description of consciousness.
- A conscious state (a quale) is *c-described* using a high dimensional mathematical structure.
Traditional Scientific Discovery
- Traditionally people have identified regularities in the physical world (Newton, Einstein, etc.).
- We generally assume that physical regularities are simple enough to be found by humans.

C-theories are Potentially Complex
- We have little or no idea about which spatiotemporal structures form CC sets.
- The mathematical relationships between c-descriptions and p-descriptions could be simple.
- Or they could be thousands of pages of differential equations.

Computational Discovery of C-theories

Methodology for Complex C-theories
- If we assume that there are simple relationships between c-descriptions and p-descriptions, we could waste a lot of time and effort looking for something that does not exist.
- Better to devise a methodology that can:
  - Identify simple relationships between c-descriptions and p-descriptions.
  - Identify complex relationships between c-descriptions and p-descriptions.
  - Prove that no relationships exist in the current data.

Computational Discovery of C-theories
- Use computers to identify relationships between c-descriptions and p-descriptions:
  - Measure consciousness and the brain.
  - Use machine learning to identify simple and complex relationships (or prove that no relationships exist).
  - This could draw on previous work in the computational discovery of scientific knowledge.
  - Could be prototyped on a simulated neural network.

APPLICATIONS
**Deductions about Consciousness**

- The science of consciousness could develop c-theories that can reliably map between p-descriptions and c-descriptions.
- Use reliable c-theories to make deductions about the consciousness of:
  - Patients in vegetative and minimally conscious states.
  - Infants.
  - Bats, octopi, snails, etc.
  - Computers and robots.

**Modification and Enhancement of Consciousness**

- A similar procedure can be used to modify and enhance our consciousness:
  1. Generate c-description of a desired state of consciousness.
  2. Use c-theory to convert c-description into p-description of a CC set.
  3. Realize the CC set in your brain.

**Modification and Enhancement of Consciousness**

- Could deduce how to repair a person's consciousness.
- Implications for implanting chips, brain uploading etc.
- The science of consciousness and the technology of brain modification have a long way to go before this will be possible.

**CONCLUSION**
Definition of Consciousness

- The physical world is invisible.
- Consciousness is a bubble of experience.
- The concept of consciousness emerged when science developed our modern interpretation of the invisible physical world.

Measurement of Consciousness

- Consciousness is measured through first-person reports.
- This raises a number of insoluble philosophical problems.
- These can be neutralized by the assumptions that I have proposed.
- The science of consciousness can be considered to be true *given these assumptions*.
- A measurement of a conscious state should be expressed as a formally structured c-description.

Measurement of the Physical World

- We cause the measured object to interact with a calibrated object.
- We observe the result and extract a number.
- The number and its corresponding object should be expressed as a p-description.

C-theories

- C-theories are a mathematical relationship between c-descriptions and p-descriptions.
- Computational methods could be used to identify c-theories automatically.
- Reliable c-theories can:
  - Deduce a system’s consciousness from a p-description of its physical state.
  - Deduce the physical state that corresponds to a desired state of consciousness.

Key Problems to be Solved

- Want an approach to consciousness that is:
  - Mathematical.
  - General.
  - Free from philosophical problems.
- The current science of consciousness lacks formats for c-description and p-description.
- If we get these right, the science of consciousness can be transformed into a data-gathering and machine learning problem.

Acknowledgements

- Many thanks to Barry Cooper and the John Templeton Foundation for supporting this work (Project ID 15619: ‘Mind, Mechanism and Mathematics: Turing Centenary Research Project’).
- I would also like to thank Anil Seth, the Sackler Centre for Consciousness Science and the Department of Informatics at the University of Sussex for hosting me as a Research Fellow during this project.
More Information

- Slides: [www.davidgamez.eu/talks](http://www.davidgamez.eu/talks)
- Some papers related to this material: [www.davidgamez.eu/publications](http://www.davidgamez.eu/publications)
- Feel free to contact me if you have any questions: david@davidgamez.eu