Consciousness and Cognition 30 (2014) 73–90

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Two theories of consciousness: Semantic pointer competition vs. information integration

Paul Thagard [a,*], Terrence C. Stewart [b]

[a] Department of Philosophy, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[b] Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, ON N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 27 November 2013

Keywords:
Binding
Competition
Consciousness
Information integration
Neural engineering framework
Neural mechanisms
Qualia
Representation
Semantic pointers

ABSTRACT

Consciousness results from three mechanisms: representation by firing patterns in neural populations, binding of representations into more complex representations called semantic pointers, and competition among semantic pointers to capture the most important aspects of an organism's current state. We contrast the semantic pointer competition (SPC) theory of consciousness with the hypothesis that consciousness is the capacity of a system to integrate information (IIT). We describe computer simulations to show that SPC surpasses IIT in providing better explanations of key aspects of consciousness: qualitative features, onset and cessation, shifts in experiences, differences in kinds across different organisms, unity and diversity, and storage and retrieval.

1. Introduction

Everyone has conscious experiences such as sensing pain, having to urinate, seeing blue, tasting chocolate, hearing music, touching wool, smelling grass, and feeling happy or sad. Consciousness also often accompanies high-level cognitive processes such as memory, learning, problem solving, decision making, and language use. Explaining consciousness is one of the most challenging problems in contemporary science, and only recently have neuroscientists dared to tackle it. The most audacious current proposal is Guilio Tononi's hypothesis that consciousness is the capacity of a system to integrate information. We will argue, however, that the information integration theory (IIT) faces serious mathematical and empirical problems.

As an alternative, we propose that consciousness is a neural process resulting from three mechanisms: representation by firing patterns in neural populations, binding of representations into more complex representations called semantic pointers, and competition among semantic pointers to capture the most important aspects of an organism's current state. Whereas IIT assumes that consciousness is a quantity that can be possessed by non-organisms such as photodiodes, countries, and the Internet, the theory of semantic pointer competition (SPC) restricts consciousness to organisms that possess sufficiently complex kinds of neural processes. We will show that SPC surpasses IIT in providing better explanations of key aspects of consciousness, including: qualitative features, onset and cessation, shifts in experiences, differences in kinds across different

* Corresponding author.
E-mail addresses: pthagard@uwaterloo.ca (P. Thagard), terry.stewart@gmail.com (T.C. Stewart).
URLs: http://cogsci.uwaterloo.ca/Biographies/pault.html (P. Thagard), http://terrystewart.ca (T.C. Stewart).

http://dx.doi.org/10.1016/j.concog.2014.07.001
1053-8100/

organisms (e.g. self consciousness in humans versus mere feeling in simpler animals), unity and diversity, and storage and retrieval.

First we give general descriptions of IIT and SPC, leaving mathematical details for appendices. Then we outline how SPC explains key phenomena about consciousness using the three mechanisms of representation, binding, and competition. We argue that these explanations are far more empirically plausible and conceptually coherent than the ones provide by IIT. In order to substantiate the claim that SPC can rigorously explain the phenomena, we present computer simulations that show that the three mechanisms can approximate the relevant results. Finally, our general discussion addresses several issues relevant to evaluating IIT and SPC, including the possibility of consciousness in non-biological systems.

2. Two theories

2.1. Information integration theory

According to IIT, consciousness is integrated information generated by a complex of elements (Balduzzi & Tononi, 2009; Koch, 2012; Tononi, 2004, 2008, 2010, 2012; Tononi & Koch, 2008). In order to ensure that we are not misinterpreting IIT, we will report it using a series of quotes from Tononi's writings.

Tononi (2010, p. 299): "Consciousness has to do with a system's capacity for information integration. In this approach, every causal mechanism capable of choosing among alternatives generates information, and information is integrated to the extent that it is generated by a system above and beyond its parts. The set of integrated informational relationships generated by a complex of mechanisms – its quale – specify both the quantity and the quality of experience. As argued below, depending on the causal structure of a system, information integration can reach a maximum value at a particular spatial and temporal grain size. It is also argued that changes in information integration reflect a system's ability to match the causal structure of the world, both on the input and the output side."

Tononi, 2010, p. 300: "High *information* means that a system's causal mechanisms can specify precisely which out of a large repertoire of potential states could have caused its current state. High *integration* means that the information generated by the system as a whole is much higher than the information generated by its parts taken independently. In other words, integrated information reflects how much information a system's mechanisms generate above and beyond its parts."

Tononi, 2012, p. 172: "Integrated information measures how much can be distinguished by the whole above and beyond its parts, and Φ is its symbol."

Tononi, 2008, p. 224: "The IIT claims that, just as the *quantity* of consciousness generated by a complex of elements is determined by the amount of integrated information it generates above and beyond its parts, the *quality* of consciousness is determined by the set of all the informational relationships its mechanisms generate. That is, *how* integrated information is generated within a complex determines not only the amount of consciousness it has, but also what kind of consciousness."

Tononi, 2008, p. 233: "If consciousness is integrated information, then integrated information exists. Moreover, according to the IIT, it exists as a fundamental quantity—as fundamental as mass, charge, or energy."

Tononi, 2008, p. 236: "The IIT implies that many entities, as long as they include some functional mechanisms that can make choices between alternatives, have some degree of consciousness."

We find Tononi's mathematical definition of Φ hard to follow, but we will attempt to clarify and assess it in Appendix A. We take his basic claim to be that consciousness is a quantity, Φ, possessed by any system (regardless of its specific causal mechanisms) that is able to generate more information (in the technical sense of reduction of uncertainty) than is generated by its parts working independently. Then consciousness in humans, other organisms, and non-biological entities is explained by their having the capacity to generate integrated information.

According to Tononi, conscious experiences arise from an organism's ability to distinguish between stimuli. Simple entities like photodiodes can only discriminate between basic features like light and dark, so their consciousness is limited. In contrast, animals can see many features and integrate them into much richer experiences resulting from choices among alternatives. We find Tononi's explanations implausible for many reasons that will be provided after we outline an alternative theory of consciousness.

2.2. Semantic pointer competition

Our theory of consciousness employs Eliasmith (2013) new idea of semantic pointers, which are representations that can function as symbols while retaining connections to sensory and motor representations. We propose that consciousness results from formation of semantic pointers and competition among them. Semantic pointers have already been useful for explaining many psychological phenomena, including recognizing patterns, serial memory, controlling motor actions, and inference (Eliasmith, 2013; Eliasmith et al., 2012), behavioral priming (Schröder & Thagard, 2013, in press); intention (Schröder, Stewart, & Thagard, 2014), emotion (Schröder & Thagard, in press), creativity (Thagard, 2014a), and concepts (Blouw, Solodkin, Thagard, & Eliasmith, 2014; Thagard, 2012). We will state the SPC theory of consciousness as a concise set of hypotheses, expound those hypotheses in more detail, and then compare SPC with IIT.

SPC, the semantic pointer competition theory of consciousness, consists of the following hypotheses: