Etica & Politica / Ethics & Politics, XIII, 2011, 2, pp. 215-234

Information and Consciousness

David Gamez
Imperial College, London
Department of Computing
david@davidgamez.eu

KEYWORDS
Philosophy of information, information integration, data integration, consciousness

1. *Introduction*

Neuroscience research often uses information measures, such as mutual information or transfer entropy, to identify the anatomical, functional and effective connections between different areas of the brain or a neural simulation (Sporns et al. 2004; Sporns 2011). Many authors have observed that integration is a key feature of conscious states (Baars 1988; Metzinger 2003), and it has been a natural progression to suggest that measures of functional and effective connectivity could be used to identify the areas of the brain that are highly integrated and thus correlated with conscious states. Tononi (2004; 2008) has gone beyond this correlation approach to claim that consciousness actually *is* integrated information, and proposed algorithms for identifying the areas of maximum information integration in a system (Tononi and Sporns 2003; Balduzzi and Tononi 2008).

The type of information integration that is claimed to be linked to consciousness has a very specific meaning because it is not just the *integration* that is important, but the *differentiation* of the information states as well. Tononi (2004; 2008) illustrates this idea of differentiated integration using the example of a digital camera sensor with a million photodiodes. This sensor is highly differentiated because it can enter $2^{1,000,000}$ different states, but in each of these states the photodiodes are acting independently and there is no integration between them. In contrast, consider a million Christmas lights connected to a single switch: when the switch is on, the lights are on; when the switch is off, the lights are off. In this system there is a high level of integration between the switch and the lights, but almost no differentiation because the system can only enter two possible states: all lights on or all lights off. In between the camera photodiodes and the Christmas lights are systems that are both differentiated and integrated: they can enter a large number of different states and these states are the result of causal interactions between the elements. According to Tononi (2004; 2008), key examples of differentiated and integrated systems are the areas associated with consciousness in the human brain.

One of the key attractions of information integration theories of consciousness is that they are precise enough to be experimentally tested. For example, Tononi's most recent theory (Tononi 2008) can predict the areas of a system that are associated with consciousness, the amount of consciousness that is present and the qualitative character of this consciousness for each state of a system. This precision of information integration theories points the way towards a more scientific approach to consciousness, in which falsifiable predictions made by different mathematically formulated theories of consciousness can be systematically compared (see Section 4). A second advantage of information integration theories is that they can be applied to both artificial and natural systems: if a link could be established between information integration and consciousness in humans, then it would be possible to make convincing predictions about the consciousness of artificial systems as well.

A major conceptual problem with information integration theories is that consciousness is a real phenomena, whereas the information states that a system holds appear to largely depend on a *subjective* interpretation - and it is difficult to see how something that is metaphysically real can be correlated with or identified with a subjective interpretation. To address this issue, the third section of this paper uses Floridi's (2009; 2010) theory of information to argue that Tononi's information integration theory of consciousness is more correctly interpreted as a *data* theory of consciousness, which has the advantage that Floridi's notion of data can be more easily linked to objective properties of a physical system. It is also suggested that Floridi's notion of information as meaningful data could be used to understand theories of consciousness that emphasize the relationship between a system and its body and environment. The last section of this paper addresses a number of philosophical and practical problems with the experimental testing of the link between data and information integration and consciousness.

2. *Background*

Research on information integration and consciousness is closely linked to work on the identification of functional and effective relationships between neurons and neuron groups using Granger causality (Granger 1969), transfer entropy (Schreiber 2000) and other measures, either using scanning or electrode data, or large-scale models of the brain. One example of this type of work is Honey et al. (2007), who used transfer entropy to study the relationship between anatomical and functional connections on a large-scale model of the macaque cortex, and demonstrated that the functional and anatomical connectivity of their model coincided on long time scales. Another example is Brovelli et al. (2004), who used Granger causality to identify the functional relationships between recordings made from different sites in two monkeys as they pressed a hand lever during the wait discrimination task. Sporns et al. (2004) and Sporns (2011) give an overview of this type of research.

A number of people have suggested that there is a link between information integration and consciousness (Baars 1988; Metzinger 2003), or that information integration actually *is* consciousness (Tononi 2004; Tononi 2008), and several algorithms for calculating information integration have been put forward. These include neural complexity (Tononi et al. 1994), stateless Φ (Tononi and Sporns 2003), state-based Φ (Balduzzi and Tononi 2008),[1] causal density (Seth 2008), liveliness (Gamez and Aleksander 2011), and an information integration measure that can be applied to time series data (Barrett and Seth 2011). Seth et al. (2006) gives a review of earlier work, identifies a number of weaknesses in Tononi and Sporns' (2003) method, and criticizes the link between information integration and consciousness.

[3] The information integration measure put forward by Tononi and Sporns (2003) is referred to as "stateless Φ" to distinguish it from the related state-based measure of Φ of Balduzzi and Tononi (2008).

There has been a limited amount of experimental work on the link between information integration and consciousness. For example, Lee et al. (2009) made multi-channel EEG recordings from eight sites in conscious and unconscious subjects and constructed a covariance matrix of the recordings on each frequency band that was used to identify the complexes within the 8 node network using Tononi and Sporns' (2003) algorithm. This experiment found that the information integration capacity of the network in the gamma band was significantly higher when subjects were conscious. Massimini et al. (2009; 2010) carried out experiments in which a TMS pulse was applied to the subject's brain and the resulting activity was recorded using EEG. Massimini et al. found that the activity resulting from the TMS pulse was more localized and less differentiated when the subjects were unconscious, suggesting that a combination of integration and differentiation is linked to conscious states. In the simulation work on information integration detailed predictions have been made about the amount and distribution of consciousness in an 18,000 neuron network (Gamez 2010), and some preliminary comparisons have been carried out between state-based Φ and liveliness (Gamez and Aleksander 2011).

3. *The Nature of Information*

"Information is notorious for coming in many forms and having many meanings. It can be associated with several explanations, depending on the perspective adopted and the requirements and desiderata one has in mind." Floridi (2010), p.1

In the work using information theory to measure brain connectivity, the exact nature of information is not important because Shannon's information equations are used as mathematical tools to identify the functional and effective connections between groups of neurons. However, the nature of information does become important when a property of the system called information integration is linked to consciousness. In this shift it is no longer the *relationships* between biological neurons that are important, but the *presence* of information integration in the system. It then becomes necessary to say what information is and how it can be identified in an arbitrary system.

To understand the problem of identifying information, imagine that an aubergine is lying on the table in front of you. Simply through its existence on the table, this aubergine contains the information that there is an aubergine on the table in front of you. Cut the aubergine open and a pattern of seeds is revealed, which can be interpreted as letters in a particular language. The genetic code of the aubergine can also be read and the aubergine contains an enormous amount of information about the location of each of its atoms relative to a particular reference point. Each of these types of information at different levels of the aubergine can be transformed into other types of information. For example, the sequence of nucleotides in the aubergine's DNA could be remapped into a sequence of numbers representing an image or sound.

This example of the aubergine suggests that physical objects can be interpreted as containing a virtually infinite quantity of information, which is highly dependently on the approach that the observer takes to the system. One theory of information that can handle this apparent relativity is the General Definition of Information (GDI). According to Floridi's (2009) formulation of the GDI, σ is an instance of information, understood as semantic content, if and only if:

GDI.1) σ consists of n data, for $n >= 1$;

GDI.2) the data are well-formed;

GDI.3) the well formed data are meaningful.

This GDI is based on the notion of *data*, which Floridi describes as a lack of uniformity in the world. The data that is accessible to us and can be read depends on pre-theoretical differences in the physical world, called *dedomena*, which cannot be experienced without an interpretation that is applied to the world. These *dedomena* are the conditions of possibility for experienced data - something like Kant's *noumena* or Locke's *substance* - that make the differences we can measure and manipulate possible. For example, *dedomena* might make the measureable difference between higher and lower charge in a battery possible, and this type of measureable difference between physical states can in turn be used to create higher levels of data, such as symbols. A definition of the scope and type of data that is available in a system is called a *level of abstraction* by Floridi. Within the brain, different levels of abstraction include the voltages of neurons' membranes, the average neuron group activity and the location and state of the molecules, atoms and ions. Experimental issues connected with the selection of a level of abstraction are discussed in Section 4.3.

From the point of view of information-based theories of consciousness, this distinction between *dedomena* and data that we understand and manipulate is important because only *dedomena* can be considered to be an objective property of the system. The data or sets of differences that we actually extract will always be the result of a particular level of abstraction that defines the scope and type of the available data. This is not a problem if we are only interested in correlations between information integration and consciousness, but it is a significant issue for Tononi's claim that information integration *is* consciousness because consciousness is typically thought to be an objective feature of the world, whereas the data that we extract from a system results from the choice of a particular level of an abstraction by an observer. One way of addressing this problem would be to use a level of abstraction that is well defined by science with precise measurement procedures, such as data at the level of fundamental particles. It would then be possible to interpret other measurements of the system, such as neuron firing activity, in the light of this underlying fundamental data. Another potential solution would be to interpret data measurements at different levels of abstraction as different levels of approximation to the underlying *dedomena*. However, a major problem would be encountered if there were radically different amounts of information integration at different levels of the system – for example, if information integration at the level of neurons was significantly different from information integration at the level of neuron groups – a problem that is discussed in Section 4.3.

Once a method for identifying data in a system has been defined, the next stage is to specify a syntax that will enable *well formed* data to be extracted in a systematic way - for example, the sequences of nucleotides that constitute the aubergine's genetic code can only be read when we can distinguish between sequences coding proteins and junk DNA. Within the brain, Floridi's notion of syntax corresponds to the different experimental techniques that are used to extract data at different levels of abstraction. For example, neurons' firing can be read by inserting electrodes and recording the voltage within a certain frequency range, and the activity of large neuron groups can be read using fMRI scanning: in each case the application of a tightly specified experimental syntax leads to well formed data, structured as voltage spikes or the levels of activation in fMRI voxels.

Although it is relatively easy to extract well formed data at a particular level of abstraction from a system, this data might not be *meaningful* in any way. Suppose that an aubergine is cut in half, covered with a grid of millimetre squares and a 1 is read off if the square contains an even number of seeds and a 0 is read off if the square contains an odd number of seeds. This sequence of 1s and 0s is well formed data because it conforms to the specified syntax, but since it lacks meaning, it is not information according to the GDI. The question of what makes data meaningful is much more difficult than the identification of data in a system, and one of Floridi's (2009) suggestions is that semantic content could be described as a combination of data and queries. So, for example, the proposition "The earth only has one moon" can be interpreted as a piece of meaningful data in which the semantic content is the question "Does the Earth only have one moon?" and the answer "yes" is a single bit of data.

Starting with the work of Shannon (1948), there has been an extensive amount of work on the communication of information, describing the entropy of an information source, the mutual information between two devices and the maximum rate of communication over a channel. However, as Floridi points out, Shannon's mathematical theory of communication (MTC) is a theory about data transmission, not information transmission, because it does not take the meaning of the messages into account: "since MTC is a theory of information without meaning (not in the sense of meaningless, but in the sense of not yet meaningful), and since we have seen that [information – meaning = data], 'mathematical theory of data communication' is a far more appropriate description of this branch of probability theory than 'information theory'." (Floridi 2009, p. 33). This suggests that "information integration" algorithms based on Shannon's work, such as Tononi and Sporns (2003), are actually measures of *data* integration, unless it can be shown that the integrated data carries semantic content.

This distinction between data and information poses an interesting question about whether there could be a link between *information* integration and consciousness, where information is understood as meaningful data. Within the brain, this type of meaningful data could be states that co-vary with the world, such as neurons in the cat visual cortex that alter their firing in response to bars of light moving in particular directions (Hubel and Wiesel 1959), or neurons that respond to the written name or picture of a famous person (Quiroga et al. 2005). Externalist theories of consciousness, such as Clark (2008) or O'Regan and Nöe (2001), might claim that there is a link between consciousness and *information* integration, understood as integration between brain states that co-vary with the world, whereas a purely brain-based theory would claim that any kind of data integration within the brain would be correlated with or *be* conscious states.

In the rest of this paper "data integration" will be used to refer to analyses of a system that do not take the meaning of the data at a particular level of abstraction into account, such as the algorithm proposed by Balduzzi and Tononi (2008), and "information integration" will be used to refer to analyses based on meaningful data. Quotation marks around "information integration" will be used to indicate the original use of the term.

4. *Testing the Link Between Data Integration and Consciousness*

4.1. *Introduction*

Tononi's (2004; 2008) "information integration" theory of consciousness (more correctly described as a *data* integration theory of consciousness) is an empirical theory about a link between a measured feature of the physical world and phenomenal experience whose strength is that it makes strong claims about the world that can be shown to be false. Many other theories of consciousness, such as higher order thought (Rosenthal 1986), might be thought to be intuitively plausible, but they are not scientific if they cannot be experimentally tested.

Experiments on the link between data integration and consciousness are likely to involve the following steps:

1. Select a system that is known or commonly agreed to be conscious.
2. Measure the data integration of the system.
3. Measure the consciousness of the system.
4. Identify correlations between data integration and consciousness.

5. Test predictions made by data integration about the consciousness of the system.

These stages are covered in sections 4.2-4.6, and Section 4.7 looks at how the data integration approach could be developed into experiments on the relationship between information integration and consciousness. While this discussion is framed in terms of Tononi's theory, a similar methodology could be applied to any theory of consciousness that is expressed in a precise mathematical or algorithmic form.

4.2. *The Platinum Standard System*

To establish whether data integration in a physical system is linked to conscious states it is necessary to start with a physical system that is known or commonly agreed to be associated with consciousness. Although we typically assume that infants and higher mammals are conscious, the only system that is confidently associated with consciousness is the awake normal adult human brain. By 'normal' it is meant that the brain is undamaged and its functions and measurements fall within two standard deviations for the human species. 'Awake' is intended in a non-technical sense to indicate that the brain is functioning in a way that is typically considered 'conscious'. This type of wakefulness is distinct from the medical definition, since apparently wakeful states can be exhibited by people in a vegetative state who are unlikely to be conscious (Laureys et al. 2002). While there will be times when the awake normal adult human brain is not conscious – for example, epileptic automatism (Ramachandran and Blakeslee 1998) - a science of consciousness has to start somewhere, and the awake normal adult human brain is the physical system that we are most certain is typically associated with conscious states.

The awake normal adult human brain will be referred to as the platinum standard system. Just as a platinum-iridium bar in Paris was used to define the length of a metre, the awake normal adult human brain is our platinum standard for a conscious system. If this physical system is not associated with conscious states most of the time, then nothing is. A further assumption, that the consciousness associated with different platinum standard systems is roughly the same, may become necessary for detailed predictions about the contents of consciousness.

It is important to note that artificial systems cannot be used to test the link between data integration and consciousness because it is not clear whether they are associated with conscious states. If a link between data integration and consciousness could be established, then it would become possible to make predictions about the consciousness of artificial systems using data integration, but this link has to be demonstrated on a platinum standard system first.

4.3. *Measuring Data Integration in the Platinum Standard System*

To investigate the link between data integration and consciousness it is necessary to measure the amount of data integration in the platinum standard system. As discussed in Section 3, only the *dedomena*, or pre-theoretical differences in the physical world, can be considered to be objectively present in the system, whereas the data that we actually extract is always the result of a choice of a level of abstraction by an observer. Some of this subjectivity can be eliminated by grounding the levels of abstraction defined by neuroscience (neuron firing,[2] activity of neuron groups) in the levels of abstraction defined by physics and chemistry (sub-atomic, atomic, molecular). However, this leaves the problem that it is far from clear whether the amount and distribution of data integration at different levels will coincide or significantly differ. As an example, consider the problem of measuring colour in a sack of oranges. If colour is measured at the level of individual oranges, then the sack of oranges will be pronounced orange. Likewise, an analysis at the level of segments will result in an orange colour, but analyses at the sub-atomic level or at the level of pips will result in zero orange colour. With data integration analyses, the key experimental challenge is to identify whether different levels coincide or contradict – for example, whether data integration analyses at the level of ions match analyses based on areas of the brain. If the data integration at different levels

[2] At the level of neuron firing, there is also the question about whether the interpretation of the neural code affects the amount of data integration – for example, rate-based analyses could give very different results from analyses based on polychromous groups (Izhikevich 2006).

does not match, then a correlations-based interpretation of the link between data integration and consciousness could make it an empirical question about which level of analysis is the most strongly correlated with consciousness. A mismatch between levels would be more problematic for an identity claim between data integration and consciousness – it might be necessary to supplement Tononi's (2004; 2008) theory with the claim, for example, that only data integration at the level of neurons *is* consciousness, which would greatly reduce its generality.

The presence of multiple data integration algorithms that apparently measure the same objective property of the physical world raises the question about which is the most *accurate* algorithm and how this accuracy can be measured. The accuracy of data integration algorithms could be evaluated by making the (problematic) assumption that data integration is correlated with consciousness, and carrying out experiments – for example, using fMRI or EEG – that measure the correlation between the output of the data integration algorithms and the reports of conscious states from the platinum standard system. If data integration is correlated with consciousness, then the algorithms that most accurately predict consciousness would be the most accurate measures of data integration. The main problem with this approach is that some or all of the algorithms might be measuring a property of the brain that is correlated with consciousness, but which has nothing to do with data integration. There is also the issue that our spatial and temporal access to the brain is severely limited, which makes it very difficult to measure data integration in humans.

Another way of measuring the accuracy of data integration algorithms is to create simulated networks with regions that we expect to have high data integration – for example, a network with several highly intra-connected modules would be expected to have greater data integration within each module. Different data integration algorithms could be run on these networks and their output compared with the areas of expected maximum data integration. This approach has the problem that our intuitions about the areas of maximum data integration might not be correct, but there does not appear to be a way of measuring the data integration of a network that does not depend on a particular algorithm. While the simulated networks approach is problematic, until our access to the brain improves it appears to be the only method available for the comparison of different measures of data integration. A set of test networks for benchmarking the accuracy of different "information integration" algorithms has been proposed by Gamez and Aleksander (2011).

A second problem with measuring data integration is the performance of some of the current algorithms - for example, it has been predicted to take 10^{9000} years to fully analyze an 18,000 neuron network using Tononi and Sporns' algorithm (Gamez 2008) and it could take a billion years to analyze a network of 30 elements using Balduzzi and Tononi's algorithm (Gamez and Aleksander 2011). Some work on addressing this issue has been carried out by Gamez and Aleksander (2011), who developed an algorithm for measuring "information integration" that scales linearly with the number of neurons and connections. However, even with these improvements, current supercomputers are likely to struggle with a full analysis of the platinum standard system, which has around 10^{10} neurons and 10^{15} synapses. Although artificial systems cannot be used as platinum standard systems, they are an ideal test environment for benchmarking the performance and accuracy of different ways of measuring data integration.

A further problem with some of the "information integration" algorithms, such as state-based Φ and liveliness, is that they rely on knowledge about the underlying causal structure of the system. This is not a problem with artificial systems, where the causal structure is usually known, but typical measurements of the platinum standard system, such as fMRI, EEG or electrode data, are sequences of states whose causal relationships are unknown. This is not an issue for Seth's Granger causality measure (Seth 2008), and it could be partially addressed in the other measures by inferring the causal structure from the sequences of states using Granger causality, transfer entropy or mutual information.

4.4. *Measuring Consciousness in the Platinum Standard System*

In experiments on the link between data integration and consciousness, the data integration of the physical system is compared with measurements of conscious states. In the platinum standard system consciousness is typically measured through first person reports, although other behaviours can be used to infer the presence of consciousness and its contents. It is also possible to use cognitive abilities that are systematically linked to consciousness to reliably infer the presence of consciousness in a platinum standard system (Shanahan 2010). There are numerous problems with the measurement of consciousness, such as the dependence on potentially fallible memory and the limited bandwidth and accuracy of human language – see Shanahan (2010) for a more detailed discussion.

4.5. *Correlations between Data Integration and Consciousness in the Platinum Standard System*

While falsifiable predictions are the gold standard for scientific theories (see Section 4.6), initial work on the link between data integration and consciousness is likely to focus on the identification of correlations between data integration and consciousness. Early experiments, such as the ones described in Section 2, are studying whether the presence of consciousness is correlated with higher data integration; eventually research will move on to examine whether different degrees of consciousness are linked to different amounts of data integration and whether the contents of consciousness vary in the way suggested by "information integration" theories.

4.6. *Predictions about Consciousness in the Platinum Standard System*

"...the real test of a scientific theory of consciousness is its ability to make falsifiable predictions: I shall certainly admit a system as empirical or scientific only if it is capable of being tested by experience. These considerations suggest that not the verifiability but the falsifiability of a system is to be taken as a criterion of demarcation ... I shall require that its logical form shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical system to be refuted by experience." Popper (2002), p.18.

A substantial amount of the current work on consciousness is based on theories that are felt to be more or less intuitively plausible. To become truly scientific, the study of consciousness has to move towards a situation in which predictions made about the consciousness of a platinum standard system are compared with the platinum standard system's behavioural reports about its consciousness.

The "information integration" theory that is most capable of making predictions (Tononi 2008) can predict the areas of a physical system that are associated with consciousness, the amount of consciousness in these areas and the qualitative character of this consciousness. When we can measure the human brain at a higher spatial and temporal resolution (and if the performance of current algorithms can be improved), it should become possible to use data integration algorithms to make predictions about the consciousness of a platinum standard system and compare these predictions with first person reports. The "information integration" theory of consciousness would become widely accepted if it could make large numbers of accurate predictions about the contents of consciousness of human subjects using only physical information about the system.

4.7. *Information Integration and Consciousness*

A possible link between information integration (integration between meaningful data) and consciousness is supported by the observation that the brain contains a large amount of unconscious data processing - regulating the body, low level motor control, etc. - whereas conscious states are mostly or entirely *about* features of our bodies and world. When I look around me I see mountains topped with snow, feel wind on my face and experience hunger in my stomach, but I am completely unaware of brain states that are unrelated to states of my body and environment.[3] This suggests that integration between meaningful data (or information, according to Floridi) might be more important for consciousness than data integration alone – a suggestion supported by the growing trend in recent years towards embodied theories of consciousness, which emphasize the links between brain, body and world (O'Regan and Noe 2001; Clark 2008). To fit in with these phenomenological observations, data integration theories of consciousness must be able to show

[3] One candidate for a non-representational phenomenal state was put forward by Block (1995), who claimed that the phenomenal content of orgasm is non-representational. This is not a particularly good example because the phenomenal content of orgasm can readily be interpreted as a representation of the internal states of a person's body, genitalia and emotion system.

that meaningful data in the brain has the highest integration, or find a different way of explaining why human phenomenal experience appears to be almost entirely about our body and world.

Analyses for data integration start with a set of connected elements, interpret their states as data, and use an algorithm to identify the balance between differentiation and integration in this data. This approach could also be used to analyze a system for information integration if the meaningful data was identified first, and then the algorithm was run on the meaningful data to measure its differentiation and integration. This could be carried out in the brain by defining meaningful data as data that *co-varies* with the world, which would be classified as *representational* data by some interpretations of this term.

Representational neurons can be identified in the brain by inserting electrodes and exposing the subject to different stimuli: if a neuron alters its firing in response to a stimulus, then it is said to be representing that stimulus. Classic work in this area was carried out by Hubel and Wiesel (1959), who inserted electrodes into the brains of cats and identified neurons that altered their activity in response to moving bars of light. More recently Quiroga et al. (2005) identified neurons in human subjects that varied their firing in response to the presence of a famous person, independently of whether the person was presented as a written word or picture. At the level of neuron groups, there has been a substantial amount of work on the identification of correlations between fMRI activity and representational content in subjects' minds. The general procedure is to measure the state of the brain using fMRI scanning, establish correlations between the fMRI patterns and features of the stimuli that the subjects are exposed to, and use the correlations to make predictions about novel stimuli. Recent work has shown that it is possible to identify patterns of fMRI activity that are linked to the images subjects are looking at (Kay et al. 2008), the words subjects are reading (Mitchell et al. 2008), and the tasks people are intending to perform (Haynes et al. 2007). As our ability to access the brain improves, it might also be possible to apply techniques that have been developed for identifying representational states in artificial systems to the brain, such as backtracking (Krichmar et al. 2005) or a combination of noise injection and mutual information (Gamez 2008).

Analyses for information integration that apply data integration algorithms to representational states will face the same performance and accuracy problems as the work on data integration and consciousness. In addition it is likely to be very difficult to exhaustively probe a system for its representational states, particularly for more abstract ones that cross multiple sensory modalities.

4.8. *The Path Ahead*

The problems identified in this paper currently make it impractical to systematically test data or information integration algorithms on the platinum standard system. Instead, initial work in this area is likely to use artificial neural networks, possibly embodied in robots, to develop more efficient algorithms and investigate novel ways of analyzing networks for data and information integration. This work on artificial systems will feed into experimental work using scanning data with low spatial and/or temporal resolution, such as the experiments discussed in Section 2. Eventually it is hoped that improvements in the speed of computers, increased spatial and temporal accuracy of brain scanning, and greater efficiency of data and information integration algorithms will make it possible to establish whether there are systematic correlations between data or information integration and consciousness in the platinum standard system.

If empirical evidence could establish that data or information integration is systematically linked with consciousness in the platinum standard system, then data or information integration algorithms could be used to make believable predictions about the consciousness of other systems, such as artificial systems, infants and animals. However, it is also possible that empirical research will demonstrate that data and/or information integration are *not* correlated with consciousness in the platinum standard system, or that they fail to make accurate predictions about consciousness. In this case, data and/or information integration theories should be abandoned and better approaches sought.

5. *Conclusions*

This paper has outlined the "information integration" theory of consciousness and used Floridi's work to provide a systematic interpretation of information that can explain how consciousness could be linked to data or information integration in the physical system. According to Floridi's definition of information, Tononi's "information integration" theory of consciousness is much more accurately described as a data integration theory of consciousness, and there is a separate question about whether there is a connection between consciousness and information, understood as meaningful data.

An overview has also been given of some of the experimental issues surrounding work on the possible link between data or information integration and consciousness. If a link between data or information integration and consciousness could be established, then it would become possible to use data or information integration algorithms to make believable predictions about the consciousness of humans, animals and artificial systems. Although "information integration" theories of consciousness are promising, this work is at an early stage of development, and much more experimental research is needed in this area.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge England ; New York, Cambridge University Press.

Balduzzi, D. and G. Tononi (2008). Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework. *PLoS Comput Biol* 4(6): e1000091.

Barrett, A. B. and A. K. Seth (2011). Practical Measures of Integrated Information for Time-Series Data. *PLoS Comput Biol* 7(1): e1001052.

Block, N. (1995). On a Confusion About the Function of Consciousness. *Behavioral and Brain Sciences* 18(2): 227-247.

Brovelli, A., M. Ding, et al. (2004). Beta Oscillations in a Large-Scale Sensorimotor Cortical Network: Directional Influences Revealed by Granger Causality. *Proc Natl Acad Sci U S A* 101(26): 9849-9854.

Clark, A. (2008). *Supersizing the Mind : Embodiment, Action, and Cognitive Extension*. New York ; Oxford, Oxford University Press.

Ferrarelli, F., M. Massimini, et al. (2010). Breakdown in Cortical Effective Connectivity During Midazolam-Induced Loss of Consciousness. *Proc Natl Acad Sci U S A* 107(6): 2681-2686.

Floridi, L. (2009). Philosophical Conceptions of Information. *Lecture Notes in Computer Science* 5363: 13-53.

Floridi, L. (2010). *Information : A Very Short Introduction*. Oxford, Oxford University Press.
Gamez, D. (2008). *The Development and Analysis of Conscious Machines*. Unpublished PhD Thesis, University of Essex, UK.
Gamez, D. (2010). Information Integration Based Predictions About the Conscious States of a Spiking Neural Network. *Consciousness and Cognition* 19(1): 294-310.
Gamez, D. and I. Aleksander (2011). Accuracy and Performance of the State-Based Φ and Liveliness Measures of Information Integration. *Consciousness and Cognition* doi:10.1016/j.concog.2011.05.016.
Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37: 424-438.
Haynes, J. D., K. Sakai, et al. (2007). Reading Hidden Intentions in the Human Brain. *Curr Biol* 17(4): 323-328.
Honey, C. J., R. Kotter, et al. (2007). Network Structure of Cerebral Cortex Shapes Functional Connectivity on Multiple Time Scales. *Proc Natl Acad Sci U S A* 104(24): 10240-10245.
Hubel, D. H. and T. N. Wiesel (1959). Receptive Fields of Single Neurones in the Cat's Striate Cortex. *J Physiol* 148: 574-591.
Izhikevich, E. M. (2006). Polychronization: Computation with Spikes. *Neural Comput* 18(2): 245-282.
Kay, K. N., T. Naselaris, et al. (2008). Identifying Natural Images from Human Brain Activity. *Nature* 452(7185): 352-355.
Krichmar, J. L., D. A. Nitz, et al. (2005). Characterizing Functional Hippocampal Pathways in a Brain-Based Device as It Solves a Spatial Memory Task. *Proc Natl Acad Sci U S A* 102(6): 2111-2116.
Laureys, S., S. Antoine, et al. (2002). Brain Function in the Vegetative State. *Acta Neurol Belg* 102(4): 177-185.
Lee, U., G. A. Mashour, et al. (2009). Propofol Induction Reduces the Capacity for Neural Information Integration: Implications for the Mechanism of Consciousness and General Anesthesia. *Conscious Cogn* 18(1): 56-64.
Massimini, M., M. Boly, et al. (2009). A Perturbational Approach for Evaluating the Brain's Capacity for Consciousness. *Prog Brain Res* 177: 201-214.

Metzinger, T. (2003). *Being No One : The Self-Model Theory of Subjectivity*. Cambridge, Mass., MIT Press.
Mitchell, T. M., S. V. Shinkareva, et al. (2008). Predicting Human Brain Activity Associated with the Meanings of Nouns. *Science* 320(5880): 1191-1195.
O'Regan, J. K. and A. Noe (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behav Brain Sci* 24(5): 939-973; discussion 973-1031.
Popper, K. R. (2002). *The Logic of Scientific Discovery*. London, Routledge.
Quiroga, R. Q., L. Reddy, et al. (2005). Invariant Visual Representation by Single Neurons in the Human Brain. *Nature* 435(7045): 1102-1107.
Ramachandran, V. S. and S. Blakeslee (1998). *Phantoms in the Brain : Probing the Mysteries of the Human Mind*. New York, William Morrow.
Rosenthal, D. M. (1986). Two Concepts of Consciousness. *Philosophical Studies* 49(3): 329-359.
Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters* 85(2): 461-464.
Seth, A. K. (2008). Causal Networks in Simulated Neural Systems. *Cogn Neurodyn* 2(1): 49-64.
Seth, A. K. (2008). Cognitive Networks in Simulated Neural Systems. *Cognitive Neurodynamics* 2(1): 49-64.
Seth, A. K., E. Izhikevich, et al. (2006). Theories and Measures of Consciousness: An Extended Framework. *Proc Natl Acad Sci U S A* 103(28): 10799-10804.
Shanahan, M. (2010). *Embodiment and the Inner Life : Cognition and Consciousness in the Space of Possible Minds*. Oxford, Oxford University Press.
Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27: 379–423, 623–656.
Sporns, O. (2011). Scholarpedia Article on Brain Connectivity. 2011, from http://www.scholarpedia.org/article/Brain_connectivity.
Sporns, O., D. R. Chialvo, et al. (2004). Organization, Development and Function of Complex Brain Networks. *Trends Cogn Sci* 8(9): 418-425.
Tononi, G. (2004). An Information Integration Theory of Consciousness. *BMC Neurosci* 5: 42.
Tononi, G. (2008). Consciousness as Integrated Information: A Provisional Manifesto. *Biol Bull* 215(3): 216-242.

Tononi, G. and O. Sporns (2003). Measuring Information Integration. *BMC Neurosci* 4: 31.

Tononi, G., O. Sporns, et al. (1994). A Measure for Brain Complexity: Relating Functional Segregation and Integration in the Nervous System. *Proc Natl Acad Sci U S A* 91(11): 5033-5037.