THEORIA 2
(2018) : 61 : str. 131-145

https://doi.org/10.2298/THEO1802131N
Originalni naučni rad
Original Scientific Paper

Janko Nešić

DOES INTEGRATED INFORMATION LACK SUBJECTIVITY?

SUMMARY: *I investigate the status of subjectivity in Integrated Information Theory. This leads me to examine if Integrated Information Theory can answer the hard problem of consciousness. On itself, Integrated Information Theory does not seem to constitute an answer to the hard problem, but could be combined with panpsychism to yield a more satisfying theory of consciousness. I will show, that even if Integrated Information Theory employs the metaphysical machinery of panpsychism, Integrated Information would still suffer from a different problem, not being able to account for the subjective character of consciousness.*

KEYWORDS: *integrated information, consciousness, subjectivity, experience, panpsychism*

Integrated Information Theory (IIT) (Tononi, 2008; Oizumi *et al.,* 2014; Tononi & Koch, 2015) is a neuroscientific theory that seeks to answer the questions about the nature of consciousness. What physical systems are conscious and of what are they conscious of? How much consciousness there is in a system if it could be measured. What is the difference in levels of consciousness in a newborn baby, an animal like a bat or a fruit fly or in vegetative patients with just „islands" of brain activity? Answering these hard philosophical and scientific questions would also have clear and significant practical merits in leading to better understanding of borderline conscious states in brain-damaged patients. The corollaries of Integrated Information seem to be that things like aggregates and machines are not conscious, though some other unexpected entities, like photodiodes, are. Still, a photodiode's experience is both quantitatively and qualitatively minimal.

In the first section of the paper, I will put forward axioms of Integrated Information Theory that are to be examined in the remainder. In the second section, I will discuss if Integrated Information Theory could answer the hard problem of consciousness, as defined by Chalmers (1996). In order to answer „the hard" question, it might be beneficial for IIT to interpret it as a version of *panpsychism*, as some philosophers have argued. This would supply the necessary metaphysics needed to answer the hard problem of consciousness. One interpretation of IIT is that it is an emergentist panpsychist theory of consciousness. This position will be examined in Section 3 of the paper. I

will argue in Section 4 that, even if IIT is interpreted as panpsychist, it would still suffer from a further problem, that it does not account for the *subjectivity* of consciousness. Versions of panpsychism that come to the aid of IIT are also plagued by the same problem. To explicate further on this drawback of panpsychist IIT I will appeal to notions of mineness and pre-reflective self-awareness that have been prominent in recent work on subjectivity in authors like Zahavi, Kriegel and Nida-Rumelin. The aim of this paper is humble: to investigate the status of subjectivity (the subjective character of consciousness) in Integrated Information Theory, that is, to examine if subjectivity is properly accommodated by IIT or to see if it is taken into account at all by this theory of consciousness.

1. Information Integration

Let us put forth some essentials of Integrated Information Theory. Adherents of IIT start by taking consciousness seriously and defend axioms and postulates concerning it. Axioms are the essential phenomenological properties of consciousness. These are taken by authors of IIT to be „immediately evident". This in itself can be very problematic, but let's take the axioms as they stand in IIT and see where they would take us. Five axioms of the Information Integration are *existence, composition, information, integration and exclusion* (Oizumi *et al.* 2014; Tononi & Koch, 2015). The first axiom, that of *intrinsic existence* (Tononi & Koch, 2015) tells us that *consciousness exists*. This is unquestionable. IIT is a theory that respects phenomenology and takes the existence of consciousness as a fact.

The second axiom, *composition*, appeals to experiential combination: „Consciousness is compositional (structured): each experience consists of multiple aspects in various combinations." (Oizumi *et al.,* 2014, p. 2). Our phenomenal field is made up of many experiences combined together. Does holism come first and out of the unity of consciousness we „carve" out the many distinct experiences later? We will come to this question later when the problem of the unity of consciousness is analyzed (Section 4).

The third axiom, *information*, claims that consciousness is *informative*: „each experience differs in its particular way from other possible experiences" (Oizumi *et al.,* 2014, p. 2). Then, with the fourth axiom, comes holism again: „Consciousness is integrated: each experience is (strongly) irreducible to non-interdependent components." (Oizumi *et al.,* 2014, p. 3). So, „experience" is also something over and above its components. Observe that this seems to go against what was said in the second axiom since the composition is nothing over and above the sum of its parts. Integration is the most important axiom because it grounds the information integration in the first place. Finally, there is the axiom of *exclusion* that states: „Consciousness is *exclusive*: each experience excludes all others – at any given time there is only one experience

having its full content, rather than a superposition of multiple partial experiences; each experience has definite borders – certain things can be experienced and others cannot; each experience has a particular spatial and temporal grain – it flows at a particular speed, and it has a certain resolution such that some distinctions are possible and finer or coarser distinctions are not" (Oizumi *et al.,* 2014, p. 3). *Exclusion* contains a *boundary* principle and an implication that no conscious subject's parts are themselves conscious.

These phenomenological truths are mirrored in *postulates* of the physical substrates that realize phenomenology. Postulates state what are the properties that physical mechanisms must have in order to support consciousness and identity is the relation between the phenomenal and the physical in IIT. Names of postulates are the same as names of axioms, but they pertain to cause-effect structures since causation as information is the „shadow" aspect of experience. For IIT, Tononi has taken Bateson's (1972) position on information: „Information is a difference that makes a difference."

Authors state in the first postulate that a system of mechanisms must exist *intrinsically*. That means that it has cause-effect power independent of extrinsic factors, „power upon itself". This is one pointer in the direction of *subjectivity* as an aspect of consciousness and towards an account of it that we find in IIT, though from a point of a system of mechanisms. Every system is structured, it has subsystems. The system has a particular cause–effect structure, given the information postulate. The system is intrinsically *irreducible* (unified) and the cause-effect structure is definite. The structure that is maximally irreducible intrinsically is a conceptual structure made of maximally irreducible cause–effect concepts, as per *Exclusion* (Tononi & Koch, 2015, p. 7).

Regarding the mechanisms that support or are the correlates of consciousness, authors say: „Integrated information is information that is generated by the whole mechanism above and beyond the information generated by its parts. This means that, with respect to information, the mechanism is irreducible." (Oizumi *et al.,* 2014, p. 7). Tononi and Koch also make the following claim of identity: „Identity: an experience is identical to a conceptual structure that is maximally irreducible intrinsically (MICS, a constellation of concepts in qualia space)... a conceptual structure completely specifies both the quantity and the quality of experience: *how much* the system exists—the quantity or level of consciousness—is measured by its Φ^{max} value—the intrinsic irreducibility of the conceptual structure; *which way* it exists—the quality or content of consciousness—is specified by the shape of the conceptual structure." (Tononi & Koch, 2015, p. 9).

Therefore, IIT should be able to measure the quantity and quality of experience in a system, the level of consciousness in a system. Even if this is true, should we conclude that IIT is a fair description of what it means to be conscious? This has to do with answering the hard problem of consciousness and, ultimately, with phenomenology and subjectivity of consciousness.

2. Problems of consciousness

Chalmers (1995) has made the canonical distinction between easy and hard problems of consciousness. The really tough problem that needs answering is why and how do we get to have consciousness in a materialistic world? How does it come about from physical systems? „How does the brain process environmental stimulation? How does it integrate information? How do we produce reports on internal states? These are important questions, but to answer them is not to solve the hard problem: Why is all this processing accompanied by an experienced inner life?" (Chalmers, 1996, p. 12). Easy problems concern the cognitive questions about consciousness. The hard problems are the ones that come from human phenomenology. The very fact that we have a phenomenology (whether it is of a feeling or a thought) is a mystery in itself. Or so it appears to be. Theories of consciousness should give an answer to the hard problem of consciousness. The question to ask now is, does IIT as a neuroscientific account of consciousness answer the hard problem?

Recently, Aaronson has criticized IIT saying that it does not solve the hard problem. He argues that it is also unable to answer „the pretty-hard" problem of consciousness. Aaronson defines „the pretty-hard" problem as a problem: „how to construct a theory that tells us which physical systems are conscious and which aren't – giving answers that agree with 'common sense' whenever the latter renders a verdict – is one of the deepest, most fascinating problems in all of science" (Aaronson, 2014). To answer this worry a theory of consciousness needs to tell us which system is conscious and this has to agree with our common sense intuitions about what conscious entities are. Since IIT posits consciousness in all sorts of counter-intuitive places and entities (like photodiodes), it does not even constitute an answer to the pretty-hard problem. Chalmers has analyzed the pretty-hard problem into several distinct problems and claims that IIT is a partial answer to at least one of them: „Construct a theory that tells us, for any given physical system, which states of consciousness are associated with that system".[1] Aaronson goes on to argue that although a large Phi value (that is central to IIT) may be a necessary condition for physical systems to be conscious, it is still not a sufficient condition for consciousness.

From the quoted Chalmers' passage, it would seem that Information Integration is a part of the answer to the *easy* questions of consciousness: „When we think and perceive, there is a whir of information-processing, but there is also a subjective aspect. As Nagel (1974) has put it, there is something it is like to be a conscious organism." (Chalmers, 1995, p. 3). It appears that IIT, after all, doesn't really approach the hard problem. IIT seems to lack the metaphysical „heavy machinery" to answer this problem. Something more is needed for IIT to constitute a valid solution to the hard problem.

1 See Mindt (2014) for discussion.

Others have already pointed out this problem with IIT and speculated that perhaps it would be beneficial to merge IIT with panpsychism (another alternative is Russellian monist panpsychism). It was pointed out that the authors of IIT already argued that the theory entails a form of panpsychism. Panpsychism would help IIT solve the hard problem of consciousness and IIT would help panpsychism solve its distinctive *problem of combination*. In this paper, I will not consider other possible interpretations of ITT, like the one that understands IIT as more akin to *the dual-aspect* theory of Chalmers (1996), but just concern myself with the panpsychist interpretation.

Since Integrated Information Theory is undefined and vague about many concepts that are of importance to the whole issue of consciousness to the extent that it may be questionable what it is a *theory of*[2], it may prove beneficial to improve on this theory with some metaphysical clarification. We need to see what could be improved in Integrated Information Theory given its merging with panpsychism. More on this will be said in the next section. I will lay out the basic postulations of panpsychism. Even though merging of panpsychism with IIT could prove important for both Integrated Information panpsychism, in Section 4 I will point out a new problem for the improved Integrated Information Theory.

3. Panpsychist Integrated Information

Roughly put, panpsychism is a metaphysical theory that claims everything in nature has a modicum of consciousness. Though it may seem counter-intuitive or even downright crazy, panpsychism was conceived in order to overcome the deficiencies and problems of both physicalism and dualism.[3] Panpsychism should successfully answer the hard problem. If IIT is interpreted as physicalist, then it would suffer the same problem of explaining consciousness as other physicalist and materialist positions. In modern literature on panpsychism it is argued that the position comes with a set of its own, new problems, those of *combination*. The most plausible form of panpsychism (the one that avoids emergence), *constitutive panpsychism* runs into difficulties.[4] Constitutive panpsychism posits that macroconsciousness is grounded in microconsciousness, macroexperience just has those microexperiences as parts and it inherits their properties. Com-

2 As Ned Block remarked at one point to Tononi: „You have a theory of something, I am just not sure what it is". See Cerullo (2015).

3 For more on this see Chalmers (2015).

4 For more on this see Coleman (2013). It was Goff (2009) who argued that set of subjects does not a priori entail the existence of another subject. Coleman goes on to strengthen Goff's claim that it is metaphysically impossible to combine subjects to yield new subjects and this leads one to dismiss constitutive panpsychism to avoid *brute* emergence.

bination problem arises when we try to understand how o-consciousness or macro-consciousness (that we pre-theoretically know in ourselves) comes from fundamental micro-consciousness of its parts. The problem is especially hard when it relates the combination of micro-subjects (if these exist) into macro-subjects of humans.

According to Chalmers, three different aspects of phenomenal states yield three different combination problems: the *subject combination* problem, the *quality combination* problem, and the *structure combination* problem (Chalmers, 2015).[5] The hardest of all problems seems to be the subject combination problem or *subject-summing*. If panpsychism is merged with IIT (or if IIT is interpreted as a form of panpsychism), perhaps, then IIT could help panpsychism solve some of these combination problems. Though it may seem possible to imagine IIT attributing to the understanding of how macro-qualities arise from micro-qualities and how experience is integrated, at least *prima facie*, it is not so easy to see how the same could hold for subject-summing. Subjects of experience don't seem to be the entities that are open to combination or processes of fusion/merging. This could even be a conceptual impossibility. Of all three problems, I will be concerned with the third, that of combination of subjects, since it is directly related to issues of subjectivity.

Is there reason to interpret IIT as a panpsychist theory? How could one corroborate such a thesis? It was Chalmers who first argued for the deep connection of information and experience: „If there is experience associated with thermostats, there is probably experience everywhere: wherever there is a causal interaction, there is information, and wherever there is information, there is experience." (Chalmers, 1996, p. 297) And a bit later: „it suggests a view in which the experiences in simple systems such as atoms are fundamental, and in which complex experiences are somehow the sum of such simpler experiences. While this is one way things could go, there is no reason that things have to go this way: complex experiences may be more autonomous than this suggests. In particular, the informational view suggests a picture on which complex experiences are determined more holistically than this" (Chalmers, 1996, p. 299). In panpsychist terms, this is close to a position like panqualityism[6], which tends to deflate the subject of experience, reducing it to just structures of qualities.[7]

---

5 Chalmers finds two additional problems: „There is the unity problem: how do microexperiences come together to yield a unified consciousness? There is the boundary problem (Rosenberg 1998): how do microexperiences come together to yield a bounded consciousness?" Roelofs (2015, 132) defines it as „Essential Boundedness (EB): The set of experiences belonging to any subject is bounded."

6 Coleman's version of panpsychism in which the basic, intrinsic properties are qualities as 'unexperienced qualia'.

7 Similarly, Cerullo (2015) has pointed out that IIT's panpsychism is a version of panexperientialism. Cerullo argues that „theories of panexperientialism, therefore, measure protoconsciousness (or proto-mentality) rather than consciousness" (Cerullo 2015, 8). So, on his interpretation, IIT would be a theory of „partial-panexperientialism". As a panpsychist, Rosenberg

One possible understanding of IIT would be to view it as a form of *emergent panpsychism*. Chalmers points to the *emergentist* inclinations of IIT: „Giulio Tononi's integrated information theory (2008), which puts forward a principle connecting degrees of integrated information with states of consciousness, can also be construed as a form of emergent panpsychism. If we see Tononi's principle as a fundamental law of nature, then it appears that macroexperiences are strongly emergent from certain physical configurations" (Chalmers, 2016, p. 16).

Luke Roelofs views Tononi as a *physicalist-panpsychist*. He points out that a physicalist combinationist like Tononi would claim that information integration is the experiential bonding[8] but that we learn of it by observing brain structure, so this will leave primitivists about consciousness unsatisfied. For Roelofs, Tononi provides a theoretical rationale for boundedness of consciousness with *Exclusion* (Roelofs, 2015, p. 132) and this is an argument against combinationism.[9] Tononi defends *Exclusion* from phenomenology and parsimony. On the other hand, Shani interprets these authors (Koch, Tononi) as saying that information integration generates consciousness, which would be a „non-panpsychist idea" (Shani, 2015). Panpsychism assumes that sentience is combined and amassed into larger unities (macro-level consciousness) form smaller ones (micro-level consciousness), not from something non-sentient. Contrary to Shani's claim, it seems to me that IIT defenders take information to be imbued with sentience on all levels (or they should if they are to avoid the danger of radical emergence, the emergence of experience from something non-experiential).

There is no aggregation in IIT (Tononi and Koch, 2015). Integration axiom implies emergence in conscious experience. Subjectivity could be what unifies experiences from inside. Also, one way to interpret *Exclusion* is as stating causal emergence in Integrated Information. I think it would be better to read *Exclusion* as stating that experience comes in discrete, unified subjects (individuals). If this integration is missing, then there is no subjectivity in IIT. It all depends on how we view subjectivity, as just relations between experiences (deflationary) or as something more. In any case, on the panpsychist interpretation (with a panpsychist „twist"), IIT is better equipped metaphysically to deal with the hard problem of consciousness. If understood as a form of panpsychism, IIT avoids postulating radical emergence. These are the good sides of panpsychism that IIT inherits. This is all well, but what about subjectivity and subject-summing? Can the panpsychist IIT avoid these problems?

(2004, Chapter 5) expounds a close form of panexperientialism, the view that experience is everywhere in nature, but it is only in a subject of experience bounded in the qualitative field. The boundaries of the field individuate subjects.

8 Phenomenal bonding relation.

9 The position of constitutive panpsychists who defend the plausibility of combination of experiences and subjects.

There are many ways one may try to solve the aforementioned combination problems. One common strategy used to answer the subject-summing problem is to *deflate* the subjects of experience, meaning that one denies that experiences need to have subjects „or at least denying that subjects are metaphysically and conceptually simple entities" (Chalmers 2015, 271). If subjects are simple entities, fusion of such subjects into larger ones seems not so plausible.[10]

Therefore, in order to alleviate the pressure of subject-summing, panpsychists deflate the subject instead. The emergence of macro-subject from micro-subjects is less plausible if subjects are fundamental, simple entities. If IIT is understood in panpsychist terms, we see the same worry expressed in the *Exclusion* postulate of Information Integration. Since IIT accommodates split-brain cases and fission of subjects, it would then seem to subscribe to the fusionist panpsychist accounts akin to that of one William Seager (Seager, 2010). Some, like Roelofs and Mørch, suggest that the phenomenal bonding relation[11] could be Information Integration. So IIT could help explain the fusion view (Seager, 2010, Mørch, 2014). Macrophenomenal properties are fusions of microphenomenal ones. Micro-subject fuse into macro-subjects at a later time (diachronic emergent combination).

It seems arguable that Integrated Information Theory would subscribe to a deflationary account of subjectivity. It would reduce subjects to experience and its relations, at best, even given its panpsychist interpretation. In the next section, I will show that IIT is lacking this important aspect that every theory of consciousness should have, that is, to be able to account for the subjective character of consciousness. Panpsycist theories that are supposed to help IIT get off the ground are also plagued by the same problem. In the paper, I will confine my criticism to Integrated Information Theory and its phenomenological axioms.[12]

---

10 For discussion, see Coleman's (2013) anti-summing arguments. Some philosophers would argue that experientiality (the quality of experience, as a broader term, as opposed to subjectivity or the „subjective character" of experience) is the constant in consciousness. Even the weaker notion of the subject is abandoned in some panpsychist accounts. This is not so hard to understand, given that experience is something that we can more easily imagine being subject to processes like fusion and fission, merging and combination. Galen Strawson's theory of SESMETs (short for „subject-of-experience-as-single-mental-thing") is not very clear on this. Although he argues for the identity of subjects and experience, and even though persistent subjects are deflated to „thin subjects", they are not eliminated from experience. Subjectivity (and with it the „thin subject") is still an aspect of an experience that has to be reckoned with.

11 Phenomenal bonding relation, a special kind of relation that holds between subjects of experience (Goff, 2009) or between experiences within a subject (this could be the „co-consciousness" relation).

12 I criticize panpsychist deflation of subjects in a different paper, Nešić (2017).

4. Intrinsic perspective

The biggest drawback of IIT and a persistent problem for the theory, even if it is improved by panpsychism, would still be that it posits no subjectivity, no point-of-view-edness. This is the claim of my paper. It would seem a conceptual truth[13] that every experience has an experiencer (subject). Any good theory of consciousness needs to explain subjectivity, say something about the nature of subjects. Even if one is not a defender of subjects as fundamental ontological entities, one has to acknowledge that something, broadly construed, as *„subjective character" of consciousness*, very likely exists.[14] So, „any satisfying theory of consciousness has to account for the first-personal access to our own consciousness" (Zahavi, 2005, p. 13). The hard problem of consciousness also demands this.

There are two distinct aspects of phenomenally conscious states, *something it's like to be* in a mental state and *what it's like to be* in that state. Phenomenally conscious mental states have a *qualitative* character and a *subjective* character (Levine, 2001; Kriegel 2009). These are separate questions of the subjectivity and quality of consciousness and these very often seem to be lumped together. The subjective designates how a certain mental state feels to the subject. On a more precise reading, it is revealed that this actually means that the subject is somehow present in experience, as a kind of self-awareness. Mineness, subjective character, for-me-ness are sometimes meant to reflect or stand in for the *pre-reflective self-consciousness* of the Phenomenological tradition. That there seems to exist something like self-consciousness in the pre-reflective and pre-conceptual sense is hold as highly plausible by many phenomenologists and philosophers of mind. It is also called *pre-reflective self-awareness*. This self-consciousness in question is not of the cognitive kind, deployed in I-thoughts, but minimal and non-reflexive. Mineness refers to the feeling that all experience in consciousness is „mine", synchronic and diachronic. But this term can have several readings and they can be very different. In her recent paper, Marie Guillot proposes that subjective character refers to several distinct notions that are being confused by some authors: *for-me-ness* (a relation of awareness between a subject and an experience), *me-ness* (a reflexive relation of awareness a subject has to itself) and *mineness* (a relation of awareness between subject and a fact that it owns the experience) and all these are about relations of awareness

13 See Chalmers (2015) for discussion.

14 Among modern philosophers who take such notions of subjectivity serious are Zahavi (2005, 2014), Gallagher (2000), Fasching (2009), Kriegel (2009), Strawson (2009), Levine (2001), Shoemaker (1996), to name just a few. *The subjective character* can thus refer to pre-reflective self-consciousness, self-awareness, mineness, etc. Nida-Rümelin (2014) shows there are three interpretations of what „subjective character" could mean: basic intentionality, primitive awareness and awareness of basic intentionality. Only in the third sense are we speaking of pre-reflexive self-awareness.

between a subject and its experiences (Guillot 2017, p. 32). The third notion is the strongest, but all three imply the presence of a subject of experience.

What is missing in IIT, and what could help us make sense of boundedness of experience, and, ultimately, make sense of Integrated Information, is the subjective character of experience. Even if IIT could help panpsychism solve the structure combination and quality combination problems, subject-summing would be left untouched. Adherents of IIT, at least, seem to be aware of the need for this subjective aspect. „Information – the ability to discriminate among a large number of alternatives – may thus be essential for consciousness. However, information always implies a point of view, and we need to be careful about what that point of view might be" (Tononi, 2008, p. 218). In IIT the maximally irreducible conceptual structure specified by a complex exists intrinsically (from its own *intrinsic perspective*). Shanahan (2015, 9) also points out that IIT: „is unable to account for the sort of self-knowledge it takes to be axiomatic".

The maximally irreducible conceptual structure is identical to the integrated experience in the same way the subject is identical to its stream of consciousness, to the total experiential field. Every experience and thus every information is always a part of some subject, always marked by the first-person givenness, so every information in a system is also marked by the subjectivity of it and there can be no other conscious subsystem (*Exclusion*).

This could mean that subjectivity is an aspect *only* of Integrated Information and of the system that realizes it. And there is the identity of phenomenological properties of experience and causal/informational properties, so it would seem that intrinsic perspective is not present in every experience as such, but *in the integrated whole* of information/experience. Authors state that maximally irreducible conceptual structure is identical to its experience (Oizumi *et al.,* 2014, p. 3). If Integrated Information is not regarded as something more than combination of experience/information, IIT is in danger of becoming a deflationist theory of consciousness. It would reduce subjects to experience or completely remove them from the metaphysical picture.

This brings us to the highly debated question of *unity of consciousness* and what it is. How should we understand the unity of consciousness? Chalmers and Bayne (2003) define it with *The Unity Thesis*: „Necessarily, any set of conscious states of a subject at a time is unified". If consciousness is unified, one needs to explain why is there such a striking unity, what brings all the conscious states into one unified experience. One way to explain the unity of consciousness is with the „co-consciousness" relation. One proposed solution to the subject-summing problem (that could be endorsed by proponents of IIT) is the *phenomenal bonding* strategy (Goff, 2009), positing a special kind of relation holding between subjects, though the nature of this relation is unknown to us. This is because we can only introspect within a subject. Chalmers (2016) proposes that the phenomenal bonding could be the „co-consciousness" relation. But what this „co-consciousness" relation really is, how to understand it?

And how is it different from phenomenal bonding? One could ask if it helps one explain the unity of consciousness or does it just state a fact.

IIT acknowledges the existence and importance of unity of consciousness as this is stated in the second and the fourth axiom. The basic claim in IIT is that consciousness (and information) is integrated, but there is no explanation as to why is that the case and how is that the case. IIT could be compatible, on the ground of *Exclusion*, with a subjectivist account of consciousness. But, the subjective character that is arguably present in our experience and is not mentioned by the current version of IIT is the first-person givenness or *mineness of experience* (Zahavi, 2005).[15] The authors of IIT claim that Information Integration is the relation that keeps the subject „bounded". They postulate the identity of consciousness and Information Integration.[16] If we say that the subject is Integrated Information, that does not include any mineness and would be in line with a deflationist theory of the self/subject.

On the phenomenological understanding, subjectivity is like a dimension or a space that grounds any phenomenal relations and experiential properties. Subjectivity could be the space where experience is manifested, and the relation of the subject to its experiences could be the same as space-time is related to its material objects. If the subjective character is conceived in the sense of mineness, it cannot be explained just by relations between experiences, it would have to refer to or imply a *subject of experience*. In the words of Wolfgang Fasching (2009): „'one awareness' (the togetherness of the manifold synchronically co-conscious experiential contents) is not a result of any relations between the experiential contents, of some synthesis of them… the character of a dimension in which the contents, with all their relations, have their presence in the first place" (2009, p. 143-144).

According to those who follow the Phenomenological tradition and accept the notions of mineness and pre-reflective self-awareness, what unifies experiences is the subject of experience, and because of it, consciousness is phenomenologically marked by *pre-reflective self-awareness*, *first-personal givenness* or *mineness* (Guillot, 2017; Zahavi, 2005, 2014; Fasching, 2009, 2011; Strawson, 2009). There is something more to a stream of consciousness than just experiences and their relations. Subjectivity thus understood implies that there might be an *experiential* or *minimal self* (Zahavi, 2014; Strawson, 2009). Dan Zahavi defends a position called *experiential minimalism* in which the for-me-ness or first-personal givenness dimension of phenomenal consciousness *is* the minimal experiential self. Though we can imagine qualitative Perfect Twins, there is a further fact that is not entailed by those qualities, and that makes them

15 Guillot argues that what Zahavi has in mind when he talks about mineness actually is for-me-ness.

16 One could imagine a possible (and rather generous) interpretation of IIT that would state that Φmax measures the value of the subjective character, the level of subjectivity and, hence, the level of consciousness of a system.

distinct: their respective individuate first-person perspectives, argues Zahavi. His position is similar to Galen Strawson's (2009) „thin" self theory[17], with the difference being that the experiential self is diachronically persistent in Zahavi's account.[18]

Saying that subjectivity is the same as Information Integration does not reveal the full truth about subjectivity. Information Integration, like phenomenal bonding, is a *relation*. If there is something like pre-reflective self-awareness in the sense that philosophers like Zahavi, Nida-Rumelin, Siewert, Guillot and others understand it, at least some routes of defending subject-summing seem to be blocked. One way to imagine how subject-summing is possible is by way of the „co-consciousness" relation (Dainton 2011). Contrary to that, I think that phenomenology of self-awareness tells us that the fact about „co-consciousness" is based in the fact that experiences belong to the same subject (not the other way round). This goes to show that regarding phenomenal bonding as „co-consciousness" is implausible.

Slors and Jongepier (2014) argue, in their paper, that mineness of experience is a product of the *external structure of experience*. They defend a reductionist, *coherentist* account of mineness and argue against Zahavi's account of mineness. As a consequence of their approach, these philosophers deny there is any phenomenological datum to mineness. There are no minimal selves in their theory. Still, they would agree with Zahavi in that „thoughts are endowed with a first-personal givenness - how can they fail to be?" (2014, p. 216). From what we have seen, this is what seems to be the case in Integrated Information as well, and a defender of Integrated Information Theory could endorse the coherentist account.

Similarly to the situation in Coleman's panqualityism, it is conceivable that a system with Integrated Information lacks awareness - though it has information/experience integration, it would be an „awareness zombie" (Chalmers) and not a subject of experience.[19] Attempts to „functionalize" awareness, and this would be the case in IIT also, eradicate the phenomenology of awareness, which is very unpalatable.[20]

---

17 Strawson's „thin" selves are brief, perishing pulses of experience, momentary subjects of experience. The human stream of consciousness is made up of many such „thin" subjects, numerically distinct thinkers and there is a special „stitching software" that holds these subjects together diachronically, bound into a single persisting subject. This leaves insurmountable gaps between short-term subject-experience-episodes.

18 Siewert pointed out that it is a category mistake to claim, as Zahavi does, that mineness (as a feature of experience) *is the experiential self*, but we can safely claim that mineness *implies* an experiential self. Mineness as Zahavi understands it, could indicate that there is something more than experiences and their relations to a subject of experience. Since integrated information is a relation it does not answer the demands of mineness.

19 See Coleman's paper in Bruntrup & Jaskolla (2016) for discussion.

20 Therefore, IIT seems to be threatened by the *nonsubject/subject gap.* See Chalmers (2016) for discussion on these issues.

5. Conclusion

There is much phenomenological evidence to the claim that something like mineness or pre-reflective self-awareness exists and plays a crucial role in our understanding of consciousness. These issues (of self-awareness) have recently come to the center of interest in modern philosophy of mind and are not to be neglected if one is to build a successful theory of consciousness. The phenomenological fact of mineness cannot be denied and dismissed and it is not accounted for in a proper way by Integrated Information Theory in its present version. Such inflationary understanding of mineness, self-awareness and subjectivity in general, creates a problem for the externalist model of the Integrated Information Theory. More work is to be done on future versions of IIT, in order to properly fit the subjective character of consciousness into the architecture of Integrated Information Theory.

Janko Nešić

References

Aaronson, S. (2014). „Why I Am Not An Integrated Information Theorist (or, the Unconscious Expander)“, *http://www.scottaaronson.com/blog/?p=1799*
Bateson, G. (1972). *Steps to an Ecology of Mind*. 1st ed. San Fransisco: Chandler Pub. Co
Cerullo M.A. (2015) „The Problem with Phi: A Critique of Integrated Information Theory“. *PloS Comput Biol* 11(9): e1004286. doi:10.1371/journal. pcbi.1004286
Chalmers, D. (1995). „Facing Up to the Problem of Consciousness”, *Journal of Consciousness Studies*, 2(3), 200-19.
Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press
Chalmers, D. (2015). „Panpsychism and panprotopsychism“ in *Consciousness in the Physical World: Perspectives on Russellian Monism*. eds. Alter, T., Nagasawa, Y., Oxford: Oxford University Press. pp. 246-277.
Chalmers, D. J. (2016). „The combination problem for panpsychism“ in L. Jaskolla and G. Brüntrup (eds.) Panpsychism, Oxford: Oxford University Press, pp. 179-214.
Coleman, S. (2013). „The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence“, *Erkenntnis* 79:19–44.
Dainton, B. (2011). „Review of *Consciousness and its Place in Nature*“, *Philosophy and Phenomenological Research* 83: pp. 238-261.
Fasching, W. (2009). „The Mineness of Experience“, *Continental Philosophy Review*, 42/2, 131–48.
Fasching, W. (2011), „“I Am of the Nature of Seeing”: Phenomenological Reflections on the Indian Notion of Witness-Consciousness“ In M. Siderits, E. Thompson, and D. Zahavi (eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*, Oxford: Oxford University Press, pp. 193–216.

Gallagher, S. (2000). „Philosophical conceptions of the self: Implications for cognitive science“, *Trends in Cognitive Sciences* 4(1): 14.
Goff, P. (2009). „Why Panpsychism Doesn't Help Us Explain Consciousness“, *Dialectica* 63:289-311.
Guillot, M. (2017). „I Me Mine: on a Confusion Concerning the Subjective Character of Experience“, *Review of Philosophy and Psychology*, 1-31.
Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. Oxford: Oxford University Press
Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford: Oxford University Press.
Mindt, G. (2014). *Integrated Information Theory and the Hard Problem of Consciousness: A New Approach to an Old Problem*. MA thesis.
Mørch, H. H. 2014. *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem. Doctoral Dissertation*, University of Oslo.
Nagel, T. (1974). „*What Is It Like to Be a Bat?*“, *The Philosophical Review*, 83(4), 435-450.
Nešić, J. (2017). „Against Deflation of the Subject”, *Philosophy and Society*, 28(4), pp. 1102- 1121
Nida-Rümelin, M. (2014). „Basic intentionality, primitive awareness, and awareness of oneself‘ in A. Reboul (Eds.), *Mind, Values and Metaphysics. Philosophical Papers dedicated to Kevin Mulligan,* Springer, pp. 261-290.
Oizumi, M., Albantakis, L. and Tononi, G. (2014). *„*From the Phenomenology of the Mechanisms of Consciousness: Integrated Information Theory 3.0”, PLoS Comput Biol, 10(5), 1-25.
Roelofs, L. (2015). *Combining Minds: A Defence of the Possibility of Experiential Combination*, Doctoral Dissertation, University of Toronto.
Rosenberg, G. (1998). „The Boundary Problem for Phenomenal Individuals“ In S. Hameroff, A. Kaszniak and A. Scott (eds.), *Toward a Science of Consciousness: The First Tucson Discussions and Debates* (Complex Adaptive Systems). Mit Press.
Schwitzgebel, E. (2014). „Why Tononi Should Think That the United States Is Conscious“. http://schwitzsplinters.blogspot.com/2012/03/why-tononi-should-think-that-united.
Seager, W. (2010). „Panpsychism, Aggregation and Combinatorial Infusion“, *Mind and Matter*, V8:2, 167-184.
Searle, J. (2013) *Consciousness: Confessions of a Romantic Reductionist by Koch, C.*, reviewed in New York Review of Books.
Shanahan, M. (2015) „Ascribing Consciousness to Artificial Intelligence”, arXiv:1504.05696
Shani, I. (2015). „Cosmopsychism: A Holistic Approach to the Metaphysics of Experience“. *Philosophical Papers* 44 (3), 389-437.
Shoemaker, S. (1996). *The first-person perspective and other essays*. Cambridge: Cambridge University Press.
Siewert, C. „Consciousness and Self-Consciousness“, Remarks on Zahavi's *Self and Others*, (PowerPoint presentation).
Slors, M. V. P., Jongepier, F. (2014) „Mineness without Minimal Selves“, *Journal of Consciousness Studies* 21 (7-8), 193-219.
Strawson, G. (2009). *Selves: An Essay in Revisionary Metaphysics*. Oxford: Oxford University Press.

Tononi, G. (2008). „Consciousness as Integrated Information: a Provisional Manifesto". *Biol. Bull.*, (215), pp. 216-242.
Tononi, G. (2012). „Integrated Information Theory of Consciousness: an Updated Account". *Archives Italiennes de Biologie*, 150, pp. 290-326.
Tononi G, Koch C. (2015) „Consciousness: here, there and everywhere?", *Phil. Trans. R. Soc. B* 370: 20140167.
Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the First-Person Perspective*, Cambridge, Mass.: MIT Press.
Zahavi, D. (2014). *Self and Other: Exploring Subjectivity, Empathy, and Shame*. Oxford: Oxford University Press.
Zahavi, D. & Kriegel, K. (2015). „For-Me-Ness: What It is and What It is Not" In D. Dahlstrom, A. Elpidorou & W. Hopp (eds.), *Philosophy of Mind and Phenomenology*. Routledge, pp. 36-53.

Janko Nešić

Da li integrisanoj informaciji nedostaje subjektivnost?
(Apstrakt)

U radu ispitujem kakav status subjektivnost ima u teoriji integrisane informacije. Postavljam pitanje da li teorija integrisane informacije rešava „teški" problem svesti. Po sebi, ova teorija ne može dati odgovor na „teški" problem svesti, ali postoji mogućnost da se teorija integrisanih informacija kombinuje sa panpsihizmom kako bi odgovorila na metafizičke probleme koji se pred nju, kao teoriju svesti, postavljaju. Argumentujem da i tako unapređena teorija integrisane informacije ne uključuje adekvatno objašnjenje subjektivnog karaktera svesti.

KLJUČNE REČI: integrisana informacija, svest, subjektivnost, iskustvo, panpsihizam