

The Philosophical Forum (2013) 44, 341-372

Theories of Consciousness as Reflexivity

Frederic Peters

Research Fellow
Psychology Department
University of New England
Armidale, NSW, Australia

Introduction

Consciousness is best understood in context, as one element of an interactive waking state in which the greater part of cognitive processing takes place in a nonconscious fashion. But if conscious and nonconscious processing are combined in the waking state, what distinguishes the former from the latter, what is consciousness, and what is its purpose? The answer to the second question depends crucially on our conclusion regarding the first. What is the property in virtue of which a state is conscious rather than nonconscious? In the following, it will be argued that of the answers most frequently proposed—

intentionality, subjectivity, accessibility, reflexivity—only the final characteristic, reflexive, auto-noetic awareness, is unique to the conscious state. Reflexivity can best be explained not as the product of a self-representational data structure, but as the expression of a recursive processing regime, in which cognition registers the properties of the processing state to a greater extent than properties of the content represented. And the principal characteristic of a reflexive processing state is cognitive reflexivity or auto-noetic awareness.

Consciousness is Reflexivity

Philosophy of mind, has over recent decades, been focused to a large extent on understanding consciousness as the expression of a particular cognitive modality—as the expression of either intentionality, accessibility, subjectivity or reflexivity. That is to say, consciousness has been equated either with (1) internal cognitive representation as such (intentionality); or (2) with the accessibility of primary informational content to various kinds of secondary processing (global broadcasting, logical thinking . . .); or (3) alternatively, on a minimalist interpretation of Nagel, with subjectivity or first-person perspective as such; or, finally and more recently (4), with reflexivity—the capacity for auto-noetic or recursive, self-monitoring awareness—as the defining characteristic of consciousness. Of these four proposed primary indices of consciousness, however, only the latter—

reflexivity—is unique to consciousness, while the other three features characterize intentional cognitive representation as such, both conscious and nonconscious.

Intentionality or representationalism holds that conscious awareness can basically be equated with representational activity as such.¹ However, as several critics have pointed out,² the assertion that conscious awareness and representational content are one and the same amounts to the claim that all intentional states are conscious as a consequence of their having intentional content, which in effect nullifies the distinction between conscious and unconscious representational states, and consequently fails as a distinguishing characteristic of the former.

Subjectivity was originally hailed as the index of consciousness by Thomas Nagel, who claimed that if conscious mentality were not configured subjectively, there would be no conscious experience, there would not be something it is like for the organism to be that organism³. Some scholars have interpreted Nagel's terse, and somewhat enigmatic language to indicate that the

¹ On consciousness as representational activity per se, see Alex Byrne, "Some Like it HOT: Consciousness and Higher-Order Thoughts," *Philosophical Studies* LXXXVI (1997): 103–129; Tim Crane, *Elements of Mind* (New York: Oxford University Press, 2001); Fred Dretske, *Naturalizing the Mind* (Cambridge MA: MIT Press, 1995); Michael. Tye, *Ten Problems of Consciousness* (Cambridge, MA: MIT Press, 1995).

² See Peter Carruthers, *Consciousness: Essays from a Higher-Order Perspective* (New York: Oxford University Press, 2005) 44-45; Robert Lurz, "Advancing the Debate Between HOT and FO Theories of Consciousness," *Journal of Philosophical Research*, XXVI (2003): 25-46, 30; Colin McGinn, "Missing the Mind: Consciousness in the Swamps, Review of Fred Dretske's *Naturalizing the Mind*," *Noûs* XXXI (1997): 528-537, 529; Amie Thomasson, "After Brentano: A One-Level Theory of Consciousness," *European Journal of Philosophy* VIII (2000): 190-209, 201.

³ Thomas Nagel, "What is it Like to be a Bat?" *The Philosophical Review* LXXXIII (1974): 435-450, 436.

first person perspective of cognitive experience, in and of itself, is sufficient for conscious awareness. Stubenberg⁴, for example, insists that the having of qualia is subjective and the subjective having of qualia (not the *knowing* that you have, just the having) is consciousness. In a similar vein, Van Gulick writes “[T]he reflexive meta-intentionality associated with conscious states . . . [derives] from the implicit self-perspectuality that is built into the intentional structure of conscious experience itself”⁵. But subjectivity, like intentionality does not discriminate conscious from nonconscious mentation⁶. Blindsight patients manually locate objects they are unaware of in relation to themselves, and nocturnal dreams retain an egocentric perspective, again without consciousness. Clearly, subjectivity characterizes cognition as such, not conscious cognition in particular.

Accessibility achieved notoriety as part of Ned Block’s claim⁷ that consciousness comes in several varieties, chief amongst which are phenomenal and access consciousness. Where the former relates to the subjective state of

⁴ Leopold Stubenberg, *Consciousness and Qualia* (Philadelphia PA: John Benjamins, 1998): 33; On the tight equation of consciousness with qualia see also John Searle, “The Problem of Consciousness,” *Consciousness & Cognition* II (1993): 310-319, 310-311.

⁵ Robert Van Gulick, “Higher-Order Global States (HOGS): An Alternative Higher-Order Model of Consciousness,” *Higher Order Theories of Consciousness: An Anthology*, ed. Rocco Gennaro (Philadelphia PA: John Benjamins) 67-92, 84-85. Making this same point, see Owen Flanagan, *Consciousness Reconsidered* (Cambridge MA: MIT Press, 1992) 194; Dan Zahavi, *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston, Ill.: Northwestern University Press) 21-22.

⁶ Joseph Neisser, “Unconscious Subjectivity,” *Psyche* XII (2006): 1-14.

⁷ Ned Block, “On a Confusion About a Function of Consciousness,” *Behavioral & Brain Sciences* XVIII (1995): 227-257. Block’s article has been cited in over 1200 publications (as of January 2013) and occasioned three additional sets of continuing commentary (1997, 1999, 2004), as well as the commentary to Block’s 2007 follow-up target article “Consciousness, Accessibility and the Mesh Between Psychology and Neuroscience,” *Behavioral & Brain Sciences* XXX, 481-548.

experiencing qualia⁸, access consciousness references, not a state, but cognitive content “poised” or available for rational analysis and use in guiding behaviour and speech⁹. Consciousness “enables” information represented in the brain to be used in reasoning, reporting and rationally guiding action¹⁰. A similar notion can be found in Tye’s styling of consciousness as abstract, non-conceptual, intentional content poised to become part of a conscious or ‘phenomenal’ state when interacting with beliefs & desires¹¹. Carruthers’ disposition HOT theory functions in much the same manner, where the conscious status of a perceptual state consists in its availability to higher-order thought¹². In the same vein, Baars’ *global broadcasting* theory indexes consciousness in terms of the global availability of information to potential use by any one of multiple processors¹³.

The guiding principle of these constructs is firmly centred on the dispositional condition (accessibility) of informational content: information that is disposed to be incorporated or configured into a conscious, as opposed to a nonconscious state, is (or can potentially become) . . . conscious. Clearly this

⁸ Block (1995): 230-231, 273-274; Tobias Schlicht, “Phenomenal Consciousness, Attention and Accessibility,” *Phenomenology and the Cognitive Sciences*, XI (2012): 309-334, 310-311.

⁹ Block (1995): 231.

¹⁰ Ibid.

¹¹ M. Tye, *Consciousness, Color, and Content* (Cambridge, MA: The MIT Press, 2000) 62.

¹² P. Carruthers, *Language, Thought and Consciousness: an Essay in Philosophical Psychology* (Cambridge: Cambridge University Press, 1996); P. Carruthers, *Phenomenal Consciousness: a Naturalistic Theory* (Cambridge: Cambridge University Press, 2000); Carruthers (2005); and see also Dan Dennett, “Toward a Cognitive Theory of Consciousness,” *Brainstorms* (Montgomery VT: Bradford, 1978) 149-17.

¹³ For the characterization of accessibility as global workspace see Bernard Baars, *A Cognitive Theory of Consciousness* (Cambridge: Cambridge University Press, 1988): 43, 74, 114; N. Block, “Concepts of Consciousness,” *Philosophy of Mind: Classical and Contemporary Readings*, ed. David Chalmers (New York: Oxford University Press, 2002) 206-218.

construct is more of a tautology than an explanation. It relates to the availability of nonconscious (or preconscious) information for inclusion within consciousness and explains little to nothing about the constitution of the conscious state itself¹⁴. Equally evident, as Weisberg, Kriegel, Burge, Schlicht and many others point out¹⁵, accessibility misconstrues consciousness the occurrent state as a mere disposition where no cognitive state (and certainly no conscious cognitive state) need actually occur at all. And finally, informational availability for processing is not associated exclusively with conscious states in any case. The great bulk of cognitive processing takes place unconsciously, and all information has to be available for processing in some sense or the system would not function. In that sense, the conditions of access could be satisfied by a nonconscious computer¹⁶. As an explanation of what distinguishes consciousness, accessibility is essentially non-functional.

¹⁴ See Jakob Hohwy, "The neural Correlates of Consciousness: New Experimental Approaches Needed?" *Consciousness and Cognition* XVIII (2009): 428-38; J.R. Searle, "Consciousness: What we still don't know," *The New York Review of Books* LII(1) 2005.

¹⁵ On accessibility as a non-event, see Schlicht (2012): 312; Josh Weisberg, "Same Old, Same Old: The Same-Order Representation Theory of Consciousness and the Division of Phenomenal Labor," *Synthese* CLX (2008): 161-181; Uriah Kriegel, "Consciousness: phenomenal consciousness, access consciousness, and scientific practice," *Handbook of philosophy of psychology and cognitive science*, ed. Paul Thagard (Amsterdam: Elsevier, 2006) 195-218; Tyler Burge, "Reflections on Two Kinds of Consciousness," *Philosophical essays, vol.II: Foundations of Mind* (Oxford: Clarendon Press, 2006) 392-419; Jennifer Church, "Fallacies or Analyses?" *Behavioral & Brain Sciences* XVIII (1995): 251-252; Thomas Natsoulas, "How Access Consciousness Might be a Kind of Consciousness," *Behavioral & Brain Sciences*, XVIII (1995): 264-265; David Rosenthal, "Phenomenal Consciousness and What It's Like," *Behavioral & Brain Sciences* XVIII (1997b): 156-157; Oliver Kauffmann, "Superblindsight, Inverse Anton, and Tweaking A-consciousness Further," *Behavioral & Brain Sciences* XXVII (2004): 290-294.

¹⁶ See Selmer Bringsjord, "Consciousness by the Lights of Logic and Commonsense," *Behavioral & Brain Sciences* XX (1997): 144-146, 145; Georges Rey, "Block's Philosophical Anosognosia," *Behavioral and Brain Sciences* XVIII (1995): 266-67, 266-267.

What then of reflexivity, of auto-noetic (self-knowing) awareness? Unlike subjectivity, intentionality and accessibility, reflexivity is unique to consciousness, and not merely unique to but the very essence of consciousness. The explicit “awareness of” or “knowing that”¹⁷ one is perceiving, thinking, feeling or behaving in particular ways is the defining characteristic of the conscious state. Informally, the centrality of the auto-noetic character is reflected in conventional linguistic usage as captured in the Oxford English Dictionary’s definition of consciousness as “the reflex act whereby I know that I think, and that my thoughts and actions are my own and not another’s.” This formal definition merely embodies a common understanding whereby it is counterintuitive to claim that conscious experiences occur without the subject being aware of them¹⁸. More to the point, it is arguable that consciousness—conscious awareness of representational content—does not occur without there being a kind of reflexive, auto-noetic awareness wherein whatever else the cognitive state might be aware of, it is intrinsically aware of its own occurrence. That conscious states do so occur is evident from the fact that they arise with awareness of their own occurrence immediately and involuntarily, that is without the need for some elaborate post hoc process of inferring or reasoning that one is

¹⁷ David Armstrong, *A Materialist Theory of Mind* (London: Routledge, 1968/1993) 227.

¹⁸ U. Kriegel, “Naturalizing Subjective Character,” *Philosophy and Phenomenological Research* LXXI (2005): 23-56, 26-27.

aware¹⁹, nor certainly the need for any secondary introspective refocusing of attention on the internal aspect of the currently cognized moment^{20,21}. And on the basis of this ongoing self-awareness, conscious states are immediately and spontaneously reportable as such, as currently occurring or as having occurred. That is, a prerequisite for being able to report that one is having (or has had) a conscious experience of this or that kind, one must be (or must have been) consciously aware of it.²²

Conversely, the principal distinguishing mark of nonconscious states is the lack of self-awareness, the absence of first-person knowability, the knowing that a particular experience is occurring. Sleep (both dreamless and dreamful), coma and anaesthetic unconsciousness are marked by a total absence of just this

¹⁹ U. Kriegel, "Consciousness and Self-Consciousness," *The Monist* LXXXVII (2004): 182-205, 198ff; D.M. Rosenthal, "A Theory of Consciousness," *The Nature of Consciousness*, eds. N. Block, O. Flanagan & G. Güzelde (Cambridge MA MIT Press, 1997) 729-754, 738.

²⁰ B. Baars, "Is Consciousness Recent?" *Consciousness and Cognition* I (1992): 139-142; O. Flanagan (1992); Arthur Reber, "The Cognitive Unconscious: An Evolutionary Perspective," *Consciousness and Cognition* I (1992): 93-103.

²¹ U. Kriegel, "Consciousness as Intransitive Self-Consciousness: Two views," *Canadian Journal of Philosophy* XXXIII (2003b): 103-132, 105; Greg Janzen, "The representational Theory of Phenomenal Character: a Phenomenological Critique," *Phenomenology and the Cognitive Sciences* V (2006): 321-339, 329; G. Janzen, *The Reflexive Nature of Consciousness* (Philadelphia, PA: John Benjamins, 2008): 106. Kriegel and Janzen enumerate four important distinctions between immediately reflexive consciousness and subsequent introspection and reflection; the former is not effortful while the latter requires deliberate effort to remain focused on just those inner mental events as opposed to external, perceptually-mediated content; the former is involuntary or automatic (you cannot choose not to be conscious) where the latter requires volition, is a matter of choice; the former is constant, ongoing, while the latter is temporary and intermittent; finally, the former is ubiquitous, self-aware at every moment where the latter is infrequent. They might have added that where consciousness includes both awareness of current perceptual input and ongoing internal commentary about that current situation—as well as past remembrances and future plans—introspective awareness refocuses attention only on the latter, internal display at the expense of perceptual awareness. Reflexivity, then, is the mechanism which establishes or generates conscious awareness of both internal and external inputs, while introspection merely focuses attention on a particular facet (internal mental information streams) of an already-conscious state.

²² Janzen (2008): 80; D.M Rosenthal, "Explaining Consciousness," *Philosophy of Mind: Classical & Contemporary Readings*, ed. D. Chalmers (New York: Oxford University Press, 2002) 406-421.

element of reflexive awareness. And the termination of these nonconscious states, the moment of waking up consists precisely in the regaining of reflexive awareness that one is in a particular cognitive state.

The claim that a state is intransitively conscious only to the extent we are transitively aware of it (or to the extent that that the state is transitively aware of itself), forms the basis of the most widely accepted definition or characterization of consciousness, Rosenthal's "transitivity principle", that consciousness is a state that I am aware of being in."²³

To say that consciousness is "a state I am aware of being in" is to claim that for the subject, this is the way it seems. However, given the frequent disparity between the way cognitive events seem and the way cognitive processes actually work—the brain does not, for example, employ round blue representations to represent a round blue ball in the mind—we need to establish

²³ D.M. Rosenthal, *Consciousness and Mind* (New York: Oxford University Press, 2005) 3-4; D.M. Rosenthal, "Thinking that One Thinks," *Consciousness: Psychological and Philosophical Essays*, eds. Martin Davies & Glyn Humphries (Oxford:Blackwell, 1993) 197-223, 199; D.M. Rosenthal, "Two Concepts of Consciousness," *Philosophical Studies* XXXIX (1986): 329–59, 335, Rosenthal (1997): 736, 742; D.M. Rosenthal, "Exaggerated Reports: Reply to Block," *Analysis* LXXI (2011): 431–437, 433; and see also Byrne (1997): 104; R.J. Gennaro, "HOT Theory Between a Rock and a Hard Place," *Journal of Consciousness Studies* XII (2005): 3-21, 13-17; R.J. Gennaro, "Between Pure Self-Referentialism and the (Extrinsic) HOT Theory of Consciousness," *Self-Representational Approaches to Consciousness*, eds. U. Kriegel & K. Williford (Cambridge, MA: MIT Press, 2006) 221–248, 222; R.J. Gennaro, *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge MA: MIT Press, 2012b) 28; Janzen (2008): 17 and Ch. 4; Kriegel (2003b): 131; Kriegel (2004):191; William Lycan, *Consciousness and Experience* (Cambridge, MA: MIT Press, 1996) 25; W. Lycan, "A Simple Argument for a Higher Order Representation Theory of Consciousness," *Analysis* LXI (2001): 3-4, 3; W. Lycan, (2004). "The Superiority of HOP to HOT," *Higher-order theories of consciousness: an anthology*, ed. R. Gennaro (Amsterdam: John Benjamins, 2004) 115-136, 93; T. Natsoulas, "The Case for Intrinsic Theory I", *Journal of Mind & Behavior* XVII (1996): 267-286, 269; David Woodruff Smith, "Rey Cogitans: The Unquestionability of Consciousness. Comment on Georges Rey," *Perspectives on Mind*, eds. H.R. Otto & J. Tuedio (Norwell: Kluwer Academic., 1988) 25-38, 28; Van Gulick (2004): 69; R. Van Gulick *Mirror, mirror—Is That All?* *Self-Representational Approaches to Consciousness*, eds. Uriah Kriegel & Kenneth Williford (Cambridge MA: MIT Press, 2006) 11-39, 12; Weisberg (2008):162, 176.

whether reflexivity is a mere subjective seeming, or self-knowing in an empirically genuine fashion. As it turns out, there is good evidence that a state's being self-aware (immediately aware of its own occurrence) is a cognitive reality. By way of general background, reflexivity can be understood as a refinement or variation of cognitive processing architecture which is built from the ground up on self-referential principles, because self-regulation of homeostasis through behaviour is the guiding principal of cognition as a whole.²⁴ That is to say, cognition extends the self-regulative homeostasis-maintaining capacity beyond the organism itself to the organism-environment interaction, by regulating self-movement in relation to homeostatic and emotional needs of the organism.²⁵

This self-regulating control of self-to-environment interaction is achieved through a cognitive system that is self-referencing in the sense of relating the outputs of various components to the needs of other components of the system. Most basically, behavioural outputs are monitored, prioritized and adjusted by homeostatic requirements for food, water, oxygen and thermoregulation²⁶, and more generally by motivational and behavioural goals.²⁷ Perceptions are also referenced against internal emotional memory before proceeding to motor

²⁴ Marcello Ghin, "What a Self Could Be," *Psyche* XI (2005): 11-39, 1-10; Pamela Lyon, "The Biogenic Approach to Cognition," *Cognitive Processing* VII (2005): 11-29.

²⁵ Patricia Churchland, *Brain-Wise: Studies in Neurophilosophy* (Cambridge, MA: MIT Press, 2002) 105; Frederic Peters, "Consciousness and Self-Regulation," *Journal of Mind & Behavior* XXX (2009): 267-290, 269-270.

²⁶ Phan Luu & Don Tucker, Self-Regulation by the Medial Frontal Cortex: Limbic Representation of Motive Set-Points, *Consciousness, Emotional Self-Regulation and the Brain*, ed. Mario Beauregard (Amsterdam: John Benjamins, 2004) 123-161.

²⁷ Karl Pribram, "A Review of Theory in Physiological Psychology," *Annual Review of Psychology* XI (1960): 1-40.

output.²⁸ More broadly, the ideomotor principle underlying perceptual control theory indicates that motor output is monitored and controlled by pre-established goals represented internally in terms of desired perceptual inputs.²⁹ Further, cognitive systems have developed an even more proactive feed-forward or anticipatory form of self-reference in the form of predictive emulation architectures that modulate output on the basis of the *anticipated* result of that output.³⁰

Self-referential architecture, particularly the internal self-adjusting loops involved in predictive cognitive activity,³¹ then, make possible the development of the kind of immediately reflexive self-cognizant processing state that constitutes consciousness. And cognitive evidence for the expression of such recursive activity in the form of genuinely reflexive self-recognizing, self-perceiving and self-knowing awareness can be gleaned from at least four distinct areas of research. At the level of personal subjective experience, consciousness arises as a single experiential field wherein distinct sensory, emotional and conceptual elements are simultaneously co-experienced as part of

²⁸ Joseph LeDoux, *The Emotional Brain* (New York: Simon & Schuster, 1996).

²⁹ Bernhard Hommel, Jochen Musseler, Gisa Ascherleben, & Wolfgang Prinz, "The Theory of Event Coding (TEC): A Framework for Perception and Action Planning," *Behavioral and Brain Sciences* XXIV (2001): 849-937.

³⁰ J.Scott Jordan, "Recasting Dewey's Critique of the Reflex-Arc Concept Via a Theory of Anticipatory Consciousness: Implications for Theories of Perception," *New Ideas in Psychology* XVI (1998): 165-187; F. Peters, "Consciousness as Recursive, Spatiotemporal Self Location," *Psychological Research PRPF* LXXIV (2010): 407-422.

³¹ Peters (2010); Daniel Wolpert, "Computational Approaches to Motor Control," *Trends in Cognitive Sciences* I (1997): 209–216.

a common state.^{32, 33} But while a unified cognitive state could be operationalized by the iterative or recurrent activation of a single schema, the resultant state would not be conscious, not self-aware, not aware of its being unified, because the mere repetition of an intentional data structure does not reverse the direction of intentionality which is antireflexive, always about something other than itself. A reflexively-processed schema on the other hand would be diachronically unified and self-knowing, aware of being so. The experience of consciousness as a consistently unified state provides strong support, then, for the contention that consciousness is genuinely reflexive in the sense of self-knowing.

Secondly, when conscious, cognition does genuinely recognize itself in the sense that it is immune to error through misidentification. One cannot think an 'I'- thought without knowing that it is in fact about oneself, because self-recognition is non-inferential, it does not rely on perceptual identification

³² Baars (1988); Tim Bayne & David Chalmers, "What is the Unity of Consciousness?" *The Unity of Consciousness: Binding, Integration and Dissociation*, ed. A. Cleeremans (New York: Oxford University Press, 2003) 23-58. Paul. Churchland, *The Engine of Reason, The Seat of the Soul* (Cambridge: MIT Press, 1995). Thomas Metzinger, "The problem of Consciousness," *Conscious Experience* (Paderborn: Schoningh, 1995) 3-40.

³³ Scholars who disagree with the notion of consciousness as a diachronic and synchronic unity, invariably hold a different notion of *consciousness* altogether, something other than subjective reflexivity. Thus Marcel envisages consciousness as introspection (Anthony Marcel, "Slippage in the unity of consciousness," *Experimental and Theoretical Studies of Consciousness*, eds. Gregory Bock and Joan Marsh (Chichester: John Wiley, 1993) 168-186, 179; O'Brien & Opie style consciousness as equivalent to the multiple perceptual streams (Gerard O'Brien & Jon Opie, "The Disunity of Consciousness," *Australasian Journal of Philosophy* LXXVI (1998): 378-95, 386-7); Semir Zeki describes self-awareness as a social experience implying the presence of others (S. Zeki, "The disunity of consciousness," *Trends in Cognitive Sciences* VII, (2003): 214-218. 217); and Dennett sees consciousness variously as cognitive content that perseveres (D.C. Dennett, *Brainchildren, Essays on Designing Minds* (Cambridge MA: MIT Press, 1998) 137), or as "a moving point of view" (Dan Dennett & Marcel Kinsbourne, "Time and the Observer: The Where and the When of Consciousness in the Brain," *Behavioral & Brain Sciences* XV (1992): 183-247, 183).

processes.³⁴ And this ongoing self-recognition has practical, empirically-observable consequences.³⁵ In Perry's (1979) illustration of following a trail of spilt sugar through supermarket aisles, only to realize that he was the careless shopper, the realization "It is I" had real psychological effects leading to immediate action (adjusting the leaky bag of sugar in his own cart). The motivational force of internal attitudes depends critically on whether the subject recognizes herself as the subject of that attitude. Consequently, self-awareness in the form of self-recognition can have a real psychological effect in terms of objectively observable behavioural expression. Consciousness can be accounted genuinely reflexive in the sense of self-recognizing.

A third source of confirmatory evidence issues from the fact that reflexivity involves a form of self-perceiving. It has always seemed self-evident, indeed logically incontestable, that when conscious, the mind is aware of itself. Thus Güzeldere notes, "The very fact of questioning the nature of my consciousness renders the fact of our not being in some way self-aware, a blatant contradiction."³⁶ The empirical reality of this self-perception is expressed in the capacity for metacognition, which requires a more basic pre-existing reflexive awareness by the mind of its own state, including the contents

³⁴ Hector Castaneda, "Self-Consciousness, Demonstrative Reference, and Self-Ascription," *Philosophical Perspectives*, ed. J.E. Tomberlin (Atascadero: Ridgeview, 1988) 405-454; Gareth Evans, *The Varieties of Reference* (Oxford: Clarendon Press, 1982); John Perry, "The Problem of the Essential Indexical," *Nous* XIII (1979): 3-21; Sydney Shoemaker, Self-Knowledge and 'inner sense'," *Philosophy and Phenomenological Research* LIV (1994): 249-314.

³⁵ Brie Gertler, *Self Knowledge* (London: Routledge) 214-215.

³⁶ Weisberg (2008): 166; G. Güzeldere, "Problems of Consciousness," *Journal of Consciousness Studies* II (1995):112-143, 115.

of that state such that I am able to know when I do or do not understand, remember or perceive such and such. Reflexive awareness then, can be accounted a genuine form of self-knowing in the form of self-perceiving.

Finally, where philosophy has concluded that self-awareness or “I-consciousness” is genuinely immune to error through misidentification, psychology provides evidence that conscious self-awareness is immune to error through misattribution—that it is not possible to seem to be awake and reflexively self-aware without actually being so. “False awakening” is conventionally described as a nonconscious, dreaming subject who thinks she has awakened when in fact she has not. This conventional interpretation appears mistaken, however, based on the false assumption that dream content only arises in nonconscious sleep states. This is not the case. Abnormal waking states such as sleep paralysis, alternate veridical perceptual content with internally-generated dream-type content³⁷, and lucid dreaming constitutes a conscious awake state where all the content is internally generated.³⁸ The presence of dream content, then is not an infallible indicator of a non-conscious sleep state. In fact, the state of “false awakening” bears all the hallmarks of an awake state wherein the subject exercises explicit metacognitive judgement (correct or not) upon her state;³⁹ remembers the content of her state;⁴⁰

³⁷ J.Allen Cheyne, *Borderlands of Consciousness: Between Dream World and Wake World* (Paper presented at Toward a Science of Consciousness, Tucson, Arizona, April 10, 2004).

³⁸ Stephen LaBerge, *Lucid Dreaming* (Boston: J.P. Tarcher, 1985).

³⁹ Giorgio Buzzi, “False Awakenings in Light of the Dream Protoconsciousness Theory: A Study in Lucid Dreamers,” *International Journal of Dream Research* IV (2011): 110-116, 114.

deliberately controls the narrative content of the dreams as it progresses;⁴¹ and remembers details of one's waking life as being of one's waking life.⁴² The fact that false awakening is in fact a genuinely awake state with dreamlike content can be taken as an indication that reflexive self-knowing cannot be simulated, that reflexivity is not a mere subjective seeming but a cognitive actuality.

Theories of Reflexivity

The principal factor differentiating conscious from nonconscious states, the primary index of consciousness as it were, can be accounted as reflexivity, immediate (pre-introspection), auto-noetic awareness which is self-knowing in the sense that it is self-recognizing, self-knowing and self-perceiving. As the very essence of what it is to be conscious, reflexivity serves as a natural basis on which to judge the viability of current theories of consciousness. Which theories do and which do not recognize reflexivity as the principle constituent of consciousness, and of those that do, which explain the mechanisms which underlie the cognitive production of reflexivity?

⁴⁰ Buzzi (2011): 69.

⁴¹ Buzzi (2011): 113.

⁴² Cheyne (2004): 9.

1. Reflexivity as Subjectivity

The most frequently encountered definition of consciousness was introduced by Nagel as being something it is like for the organism to be that organism.⁴³

Nagel's somewhat enigmatic phrasing has been interpreted by some scholars as primarily referencing the subjectivity or first-person perspective of cognitive experience, without any reference to reflexivity.⁴⁴ By contrast to this straightforward reduction of conscious reflexivity to subjectivity, a second group of scholars—Zahavi, Janzen and Kriegel—clearly identify consciousness with subjective reflexivity, but relate reflexivity to subjectivity in distinctly opposite ways. Dan Zahavi overtly casts the central constituting characteristic of consciousness as reflexivity or self-awareness, stating that

[I]n much phenomenological literature, the discussion of self-awareness is not so much a discussion of how consciousness is aware of a self . . . as it is a discussion of how consciousness is aware of itself. . . . Consciousness is self-luminous. It is characterized by intentionality, but being intentionally aware of objects, it is simultaneously self-aware through and in itself. Its self-awareness is not due to a secondary act or reflex but is a constitutive moment of the experience itself, and

⁴³ Nagel (1974):436.

⁴⁴ C. McGinn, *The Problem of Consciousness* (Oxford: Blackwell, 1991); Charles Siewert, *The Significance of Consciousness* (Princeton, NJ: Princeton University Press, 1998); Searle (1993); Siewert (1998); Stubenberg (1998); Van Gulick (2004). And see F. Peters, "Consciousness and Subjectivity," *Under Review*, for extended discussion.

consciousness can consequently be compared to a flame, which illuminates other things, and itself as well.⁴⁵

He then goes on to claim, with Flanagan, that reflexivity is inherent in or an intrinsic concomitant of subjectivity:

[I]nsofar as there is something it is like for the subject to have experiences, there must be some awareness of these experiences themselves; in short, there must be self-awareness . . . Thus, in our view, phenomenal consciousness is simply a primitive type of self-awareness, and we can therefore only agree with Flanagan when he writes: ‘. . . all subjective experience is self-conscious in the weak sense that there is something it is like for the subject to have that experience. This involves a sense that the experience is the subject’s experience, that it happens to her, occurs in her stream’ (Flanagan, 1992, p. 194).⁴⁶

Again, stressing the equivalence of “self-givenness” with “subjectivity”, Zahavi reiterates that subjectivity entails or brings with it an intrinsic form of self-referentiality or consciousness:

Mary might certainly realize that John is in pain, she might even sympathize with John, but she cannot actually feel John’s pain the same way John does. Mary has no access to the *first-personal givenness* of John’s experience. This is not something quite incidental to their being, a mere varnish that the experiences could lack

⁴⁵ D. Zahavi, “Phenomenal Consciousness and Self-Awareness: A Phenomenological Critique of Representational Theory,” *Journal of Consciousness Studies* V (1998): 687-705, 689, 696; Shaun Gallagher & Dan Zahavi, *The Phenomenological Mind* (London: Routledge, 2008) 51, where he quotes Frankfurt on consciousness as “reflexivity”. See also D. Legrand, “Pre-reflective self-as-subject from experiential and empirical perspectives,” *Consciousness & Cognition* XVI (2007): 583-599.

⁴⁶ Zahavi (1998): 690-3; D. Zahavi, *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston, Ill.: Northwestern University Press, 1999).22; D. Zahavi, *Subjectivity & Selfhood* (Cambridge, MA: MIT Press, 2005a) 46. His reference is to Flanagan (1992).

without ceasing to be experiences. On the contrary, it is this first-personal givenness that makes the experiences *subjective*. To put it differently, with a slightly risky phrasing, their first-personal givenness entails a built-in self-reference, a primitive experiential self-referentiality.⁴⁷

In linking self-awareness with subjectivity, Zahavi follows Husserl who, he says, “took self-awareness to be an essential feature of subjectivity.”⁴⁸

Like Zahavi (and Husserl), Janzen also cites Flanagan (1992, p. 194), arguing that “phenomenal character” or what-its-likeness includes or constitutes “a type of self-awareness”

Flanagan assumes, without argument, that the what-it-is-likeness or phenomenal character of our experience is a type of self-awareness. My aim is to elaborate on this suggestion, to defend the thesis that phenomenal character is constituted by a type of self-awareness, i.e. by a low level or implicit self-awareness that is built into every conscious state.⁴⁹

Zahavi and Janzen, like van Gulick, recognize that a central characteristic of consciousness is reflexive awareness, but unlike van Gulick, they do not reduce reflexive consciousness to subjectivity. Rather, they hold that subjectivity intrinsically manifests or gives rise to self-awareness, to a Rosenthalian being aware that one is in the cognitive state. This position amounts to a kind of *cognitive* panpsychism. Where panpsychism proper insists that some minimal

⁴⁷ D. Zahavi, “Intentionality and Experience,” *Synthesis Philosophica* XL (2005b):.312; cf. Zahavi (1999): 21.

⁴⁸ D. Zahavi, *Subjectivity & Selfhood* (Cambridge, MA: MIT Press, 2005a) 50.

⁴⁹ Janzen (2008): 155; and see Janzen p. 28 where what-it-is-likeness or “phenomenal character” is a criterion for state consciousness, providing its necessary and sufficient conditions.

form of cognitive capacity is a fundamental property of the physical universe,⁵⁰ *cognitive* panpsychism would insist that a minimal form of conscious self-awareness accompanies every subjective cognitive act because subjectivity entails more than simply being the cognitive state (a la Stubenberg and Searle), more than an implicit registration of the world in relation to self (as with Van Gulick). Subjective cognition, they claim, has self-awareness, at least in some minimal form, because self-awareness is simply a fundamental property of subjective cognition. As Gennaro suggests,⁵¹ this is this is presumably Nagel's position as well.

Zahavi and Janzen's claim that (minimal) self-awareness is inherent in every subjective cognitive act can be contrasted with the position of Kriegel, an intrinsic self-representational theorist, who insists on the polar opposite, that reflexivity is primary and subjectivity essentially a derivative expression of it. Kriegel writes,

It is unlikely there could be anything it is like for a subject to be in a mental state she is unaware of being in . . . [consequently] intransitive self-consciousness is a necessary condition for phenomenal consciousness: unless M is intransitively self-conscious, there is nothing it is like to be in M, and therefore M is not a phenomenally conscious state.⁵²

⁵⁰ D. Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996) Ch. 8.

⁵¹ R.J. Gennaro, *Consciousness and Self-Consciousness* (Amsterdam: John Benjamins, 1996) 15.

⁵² Kriegel (2003b): 106; cf. Kriegel (2005): 23-26; Kriegel (2006): 200; U. Kriegel, *Subjective Consciousness: a Self-Representational Theory* (New York: Oxford University Press, 2009) 104. Similar sentiments have been expressed by Rosenthal (1997):733; Rosenthal (2011): 433; W. Lycan,

Is reflexive self-awareness inherent in subjectivity, as Zahavi, Janzen and presumably Nagel propose, or is subjectivity a derivative expression of reflexivity as outlined by Kriegel?

The evidence indicates that neither of the abovementioned positions is an accurate portrayal: reflexivity is not inherent in subjectivity, nor is subjectivity merely an expression of reflexivity. Subjectivity and reflexivity are two distinct cognitive characteristics or properties such that while subjectivity characterizes all cognitive events, not all cognitive events are conscious. and therefore subjectivity is not the factor distinguishing conscious from nonconscious cognition. That is, all cognitive experience—whether conscious or unconscious—is subjective in the sense that it only exists for the subject operationalizing that cognitive state; the subject has literally to be (or be in) that cognitive state in order to realize or have access to those cognitive characteristics—no objective observer can apprehend or register another's thoughts feelings or sensations. In that sense at least, subjectivity is ontologically subjective as Searle maintains.⁵³ But not all subjectively-realized cognitive experience is necessarily conscious.

Each of the epistemic dimensions of subjectivity—privileged or immediate access, first-person perspective and reference to a self—characterizes nonconscious as well as conscious cognition. Where immediate cognitive access

“Consciousness as internal monitoring,” *AI, Connectionism and Philosophical Psychology. Philosophical Perspectives IX* (1995): 1-14, 3; and even, on occasion, Zahavi (1999): 23-4.

⁵³ J.R. Searle, *The Rediscovery of the Mind* (Cambridge, MA: MIT, 1992) 93-95.

to its contents is considered a mark of consciousness,⁵⁴ this same lack of mediation by inferential or deliberative processing is also characteristic of unconscious sensory-motor reflex arcs. Again, conscious subjectivity is frequently said to involve me-ness, an explicit ascription or awareness of a self,⁵⁵ but nonconscious dreaming is also, routinely constructed around a self structure.⁵⁶ And finally, subjectivity is characterized by an explicit first-person perspective, whereby the spatially extended manifold is configured egocentrically in relation to the subject as centerpoint, such that perceptions involve the use of what are often called egocentric directional vectors such as up, down, left right, behind, in front and so forth. But this spatial first-person perspective operates nonconsciously. Blindsight patients can manually locate objects without conscious awareness of their location vis-à-vis the subject.⁵⁷ Nocturnal dreams retain an egocentric perspective, again without

⁵⁴ Joseph Levine, "Consciousness and (Self-)Awareness," *Self-Representational Approaches to Consciousness*, eds. U. Kriegel & K. Williford (Cambridge MA: MIT Press, 2006) 173-198.

⁵⁵ Miguel Sebastian, "Experiential awareness: Do you prefer it to me?" *Consciousness Online Conference 2012*. URL= <<http://consciousnessonline.com/2012/02/17/experiential-awareness-do-you-prefer-it-to-me/>>

⁵⁶ William Farthing, *The Psychology of Consciousness* (Englewood Cliffs NJ: Prentice-Hall, 1992); Patrick McNamara, "The Appearance and Role of the Self in Dreams," *Dream Catcher: The neuroscience of our night life* (July 25, 2011). Available at <<http://www.psychologytoday.com/blog/dream-catcher/201107/the-appearance-and-role-the-self-in-dreams>>.

⁵⁷ T. Schlicht, "Non-Conceptual Content and the Subjectivity of Consciousness," *International Journal of Philosophical Studies* XIX (2011): 491-520, 504-5; Larry Weiskrantz, *Blindsight. A Case Study and Implications* (New York: Oxford University Press, 1986).

consciousness.⁵⁸ Non-spatial egocentric reference frame value judgements are exercised nonconsciously⁵⁹ as are self-relating emotions.⁶⁰

Subjectivity, then, is characteristic of cognition in general, both conscious and nonconscious.⁶¹ As Carruthers puts it,⁶² non-conscious subjectivity involves mere sensory/perceptual engagement with the world (“worldly subjectivity”), whereas consciousness involves awareness of one’s experience of the world (“experiential” subjectivity). As a basic characteristic of cognition, subjectivity is not intrinsically self-aware (as Claimed by Flanagan, Zahavi and Janzen), nor certainly, dependent upon and merely a derivative expression of self-awareness (as claimed by Kriegel). As the basic substratum of all cognition, it lacks the critical ingredient that differentiates conscious from nonconscious states.

2. Higher- Order Theories of Reflexivity

Where first order accounts of consciousness discount or ignore reflexivity altogether, effectively removing the distinction between conscious and nonconscious states, higher order (HO) representational theories are explicitly focused on explaining this distinction in terms of the presence of a secondary

⁵⁸ Farthing (1992): 259.

⁵⁹ William Kunst-Wilson & Robert Zajonc, “Affective Discrimination of Stimuli That Cannot be Recognized,” *Science* CCVII (1980): 557–558; Richard Nisbett. & Timothy Wilson, “Telling More Than We Can Know: Verbal Reports on Mental Processes,” *Psychological Review* LXXXIV (1977): 231-259.

⁶⁰ Neisser (2006): 8; M. Nussbaum, *Upheavals of Thought: The Intelligence of Emotions* (New York: Cambridge University Press, 2001) 61-4, 147.

⁶¹ Neisser (2006): 1; Kai Vogely & Gereon Fink, “Neural Correlates of First-Person Perspective,” *Trends in Cognitive Science* VII (2003): 38-42.

⁶² Carruthers (2000): Ch. 6

state directed at the primary world-directed state. Several cognitive data structures have been proposed to operationalize the crucial *awareness-of* capacity: higher order thought theory (HOT) insists that a first order state is the focus of an additional thought to the effect that the first order state is occurring.⁶³ Both higher order belief theory⁶⁴ and higher-order judgement theory⁶⁵ invoke a dispositional state in relation to first order content, while higher-order feelings theory calls upon a second-order emotional state.⁶⁶ Higher-order source monitoring⁶⁷ invokes a species of executive processing. Higher-order global state theory (HOGS) turns to an all-encompassing self-to-world representational state into which primary content is incorporated.⁶⁸ On the other hand, inner sense, or higher-order perception theories (at least those of Armstrong and Lycan), although normally included with the abovementioned

⁶³ Carruthers (2000); Farthing (1992); Genarro (1996); Natsoulas (1992); Rosenthal (1986); Rosenthal (1993); Rosenthal (1997); Shoemaker (1994); T. Natsoulas, "The concept of Consciousness3: The Awareness Meaning," *Journal for the Theory of Social Behavior* XXII (1992): 199–225; J. Weisberg, "Active, Thin, and HOT! An Actualist Response to Carruthers' Dispositionalist HOT View," *Psyche*, V/6 (1999).

⁶⁴ D. Dennett, "Who's on First? Hetero-Phenomenology Explained," *Journal of Consciousness Studies* X (2003): 19-30; D.H. Mellor, "Conscious Belief," *Proceedings of the Aristotelian Society* LXXVIII (1977-78): 87-101.

⁶⁵ D. Dennett, *Consciousness Explained* (Boston: Little, Brown and Company, 1991); J. Fodor, "Fodor's Guide to Mental Representation," *Mind* XCIV (1985): 76-100; Allison Gopnik, "How we know our Minds: The Illusion of First-Person Knowledge of Intentionality," *Behavioral & Brain Sciences* XVI (1993): 1-14; Hakwan Lau, "A Higher Order Bayesian Decision Theory of Consciousness," *Progress in Brain Research* CLXVIII (2008): 5–48; A.J. Marcel, "Phenomenal Experience and Functionalism," *Consciousness in Contemporary Science*, eds. Anthony Marcel & Eduardo Bisiach (New York: Oxford University Press, 1988) 121-158.

⁶⁶ Michael Gazzaniga, "Consciousness and the Cerebral Hemispheres," *The Cognitive Neurosciences*, ed. M.S. Gazzaniga (Cambridge MA: MIT Press, 1995) 1391-1400.

⁶⁷ Robert Kunzendorf, "Self Consciousness and Monitoring Conscious States," *Imagination, Cognition and Personality* VII (1987-88): 3-22.

⁶⁸ Van Gulick (2004); cf. T. Metzinger, "Subjectivity of Subjective Experience," *Neural Correlates of Consciousness* (Cambridge MA: MIT Press, 2000), and note 135 below.

approaches, should not be considered a higher-order construction like the others because they do not relate to the mechanism initially constructing or creating consciousness, but rather to the application of a subsequent metacognitive process—introspection--over an already-conscious state.⁶⁹ Conscious reflexivity, as noted above, is immediate and effortless, occurring without the need for deliberate post hoc inference, nor certainly any secondary introspective refocusing of attention.

The central proposition of all of the higher-order constructions is that consciousness is created as the result of a second order cognitive data structure (a thought, perception, judgement, feeling) being directed at or arising in relation to a first order cognitive event, such that the combination of these two cognitive states creates consciousness. But this claim has raised a host of objections, leading some analysts to question the basic viability of higher-order theory altogether.⁷⁰ The more significant criticisms include, firstly, that all

⁶⁹ Higher-order theories are customarily addressed in terms of a contrast between HOP and HOP; the former approach involving a second-order thought-like state, and the latter a perception-like process “scanning” first order content. On closer inspection however, it is evident, that the two leading proponents of HOP (Lycan and Armstrong) have consistently pointed to *introspection* as the cognitive process that produces consciousness (Lycan (1996): 13-14; cf. (2004): 101-102; (1987): 72; W. Lycan (1987). *Consciousness* (Cambridge, MA: MIT Press) 72; Armstrong (1997): Armstrong, D. (1997). What is consciousness. In Ned Block, Owen Flanagan & Güven Güzeldere (Eds.), *The nature of consciousness*. Cambridge MA: MIT Press) 721-728, 723-723. However introspection is better regarded as an attentional redirection process within an existing conscious state (Baars (1992); Flanagan (1992); Reber (1992); Stubenberg (1998). As pointed out in Note 21 above, pre-reflective consciousness (reflexivity) is effortless, involuntary, constant, ubiquitous and encompasses both perception and internal thought, while the latter (introspection) is effortful, deliberate (requires choice), infrequent, temporary and focused only on the internally-generated stream of cognitive content. The two, then, are quite distinct.

⁷⁰ U. Kriegel, “Consciousness, Higher-Order Content and the Individuation of Vehicles,” *Synthese* CXXXIV (2003c): 477-504. N. Block, “The Higher-Order Approach to Consciousness is Defunct,” *Analysis* LXXI (2011): 419-431.

higher-order theories of consciousness are structured in terms of an awareness of first-order content, and are thereby directly at odds with evidence to the effect that perceptual, conceptual and emotional content is neither sufficient for the production of consciousness (as evidenced by the substantial domain of nonconscious cognition), nor even necessary for consciousness (as suggested in the various forms of dissociation of informational content from awareness in cases of hemispatial neglect, blindsight, aphasia and agnosia.⁷¹ Indeed, if we consider the fact (see sec. 3 below) that cognition includes the reading of the features of what is being represented as well as of characteristics of the representing state, it becomes apparent that conscious reflexivity reflects the latter (state properties) and as such, is therefore always distinct from the ever-changing streams of informational content relating to what is being represented. On this basis, consciousness is independent of informational content: consciousness is not equivalent to an awareness-of-qualia.

But higher-order theories are also seriously challenged in their effort to satisfactorily account for the generation of the reflexive *awareness-of* element even within the overall qualia (awareness-of-content) construct. Higher order theories attempt to account for this “awareness-of” in terms of “the mind directing its intentional aim upon its own states and operations” as Zahavi puts

⁷¹ For details see F. Peters, “Consciousness is Distinct From Qualia,” *Submitted, under review* (2013b).

it.⁷² However, crucially, the higher order construct itself which is supposed to supply the self-awareness is not itself self-aware,⁷³ so that the critical “awareness of” is purportedly generated by a nonconscious thought, belief, judgement or feeling which is about something other than itself. But beliefs, judgements, feelings and thoughts about something else all can and routinely do take place nonconsciously,⁷⁴ precisely because they are not inherently self-aware and do not engender self-awareness. The non-conscious status of the higher-order construct constitutes one of the most frequently voiced criticism of higher order theories, that a basically nonconscious cognitive construct cannot imbue consciousness in another nonconscious first-order construct or process⁷⁵.

This criticism, commonly referred to as *the problem of the rock*, turns on the point that a rock doesn't become conscious simply by my looking at it, so why

⁷² D. Zahavi, “Back to Brentano?” *Journal of Consciousness Studies* XI (2004): 66–87, 67.

⁷³ Carruthers (2000): xiv-xv; Rosenthal (2005): 743.

⁷⁴ Weiskrantz (1986); Reber (1992); John Kihlstrom, “The Continuum of Consciousness,” *Consciousness & Cognition* II (1993): 334-354; Max Velmans, “Is Human Information Processing Conscious?” *Behavioral and Brain Sciences* XIV (1991): 651-726.

⁷⁵ Versions of this criticism include Byrne (1997): 110ff; Gennaro (2004):.6, (2006): 225; Janzen (2008): 93-4, 113-4; Natsoulas (1996): 276; Stubenberg (1998):185-186; Thomasson (2000): 198; N. Block, “Consciousness,” *A Companion to the Philosophy of Mind*, ed. S.D. Guttenplan (Oxford: Blackwell, 1994) 210-219, 212; P. Carruthers, “Higher-Order Theories of Consciousness,” *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition). URL = <<http://plato.stanford.edu/archives/fall2011/entries/consciousness-higher/>> ; Victor Caston, “Aristotle on Consciousness,” *Mind* CXI (2002): 751-815, 780; Paula Droege, *Caging the Beast: A Theory of Sensory Consciousness* (Philadelphia, PA: John Benjamins, 2003) 38-39; Alvin Goldman, “Consciousness, Folk Psychology and Cognitive Science,” *Consciousness & Cognition* II (1993): 364-382,366; U. Kriegel, “Consciousness, Permanent Self-Awareness and Higher-Order Monitoring,” *Dialogue* XLI (2002): 517-540, 522-523; Pessi Lyyra, “Two Senses for ‘Givenness of Consciousness’,” *Phenomenology and the Cognitive Sciences* VIII (2009): 67-87, 70; Mark Rowlands, “Consciousness and Higher-Order Thoughts,” *Mind and Language* XVI (2001): 290-310, p. 291; R. Van Gulick, “Inward and Upward: Reflection, Introspection, and Self Awareness. *Philosophical Topics*, 28, (2001):275-305, 294, (2004): 71-72, (2006): 14; D. Zahavi, Phenomenal Consciousness and Self-awareness: a Phenomenological Critique of Representational Theory,” *Journal of Consciousness Studies* V (1998): 687-705, 693, 695.

should a nonconscious cognitive data structure become conscious simply as a consequence of another data structure, or dispositional state being directed at it, (or being a representational redescription of it or simply being “about” it), particularly if the latter, second-order construct is nonconscious as well?

This problem of the rock also constitutes the main stumbling block for Gennaro’s wide intrinsicity view (WIV) theory, which, like other HOR theories, maintains that consciousness involves an awareness of being in the state one is in (2006, p.222), but unlike other theories, insists that the cognitive construct is a single, complex first-order state with two elements or parts, one directed at the world and the other (a metacognitive thought--MET) directed at the first part, thereby generating, according to Gennaro, a self-referential or conscious cognitive moment (1996, p. 28).⁷⁶ It remains to be seen, however,

⁷⁶ Gennaro’s WIV theory seeks to avoid many of the criticisms levelled at the standard two-state construct of higher-order theories (particularly Rosenthal’s HOT) by stipulating that the two elements are part of one complex state. This is similar to Brentano’s one-state “intrinsic” theory, but Gennaro also wishes to avoid what he sees as Brentano-style claims that the second, consciousness-making element is somehow reflexively directed at itself, which he insists is impossible (2006:226ff; 2012: Ch5, and see also Gennaro’s “Comments on Miguel Sebastian’s ‘Experiential Awareness: Do you prefer *It* to *Me*?,” *Online Consciousness Conference 4* (2012):3-4 where he notes similar sentiments by Kriegel). In an earlier piece (1996:28) he writes, “[s]trictly speaking, on the WIV, a CMS [conscious mental state] is not directed at itself. There is, rather, an ‘inner directedness’ or ‘inner relationality’ in the CMS. The MET is not directed back at the entire complex CMS, but rather at the [first order] psychological state it renders conscious. Conscious mental states are not about themselves, but there is a kind of indirect self-reference.” For Gennaro, then, a conscious state is “self-referential” (2006:222), not because the state as a whole is reflexively directed back upon itself, but because one element within the state is referring to or representing another element of that state. Although Gennaro’s theory contains elements of both intrinsic and higher-order approaches to explaining conscious, he identifies himself as a HOT theorist (2005: 9; R. Gennaro, “Are there Pure Conscious Events? *Revisiting Mysticism* eds. C. Charabarti & G. Haist (Cambridge: Scholars Press, 2008) 100-120, 107), and is therefore treated as such here. As he notes, however (2012: 3-4), his two-parts-within-one-state construct bears strong similarities to the mechanism proposed by Kriegel’s intrinsic theory. According to the view of this paper, both approaches miss the mark when explaining consciousness as qualia, as opposed to reflexive awareness as such; but both are heading in the right direction in their emphasis on the importance of the processing regime, rather than basing the self-referential capacity of consciousness solely on the representational content of data structures

why a relocation of this same two-part mechanism within a single state would render the dynamic (one representational element being intentionally ‘directed at’ or simply about another element) capable of generating conscious self-awareness when it is unable to do so in a two-state configuration.⁷⁷

Finally, it is basically questionable, in any case, whether higher order mental constructs (beliefs, percepts, thoughts, judgements), either extrinsic to the first-order state or intrinsic as in Gennaro’s model, actually do monitor or arise in relation to first-order states one-on-one—and in fact, it seems they do not. Consciousness is unified not only synchronically but also diachronically – a constant flow. But the evidence from introspection⁷⁸ and thought sampling studies⁷⁹ is that while inner cognitive processing (thinking, daydreaming, remembering) does proceed in a more or less unbroken stream, it is not necessarily or even generally related to or directed toward the current perceptual input. These inner judgements, thoughts, beliefs and so on proceed more or less independently of first-order perceptual content, and are therefore not related to it in the way that higher-order theory requires. Moreover, even if that higher-order content were directed toward every instance of first-order input, there is

(Gennaro 2004:62-63, 2006: 239-41; Kriegl 2009: 224, n.37, 2005: 49-51; U. Kriegel, “A Cross-Order Integration Hypothesis for the Neural Correlate of Consciousness,” *Consciousness & Cognition* XVI (2007a): 897-912, 901.

⁷⁷ Pace Gennaro (2004):7, (2005):4ff, (2006):225.

⁷⁸ Byrne (1997): 117.

⁷⁹ John Antrobus, Jerome Singer & Stanley Greenberg, “Studies in the Stream of Consciousness: Experimental Suppression of Spontaneous Cognitive Processes,” *Perceptual and Motor Skills* XXIII (1966): 399–417; Eric Klinger & W. Miles Cox, “Dimensions of Thought flow in everyday Life,” *Imagination Cognition and Personality* VII (1987–1988): 105–128.

simply no possibility, analysts point out, that cognition has a store of higher-order conceptual constructs sufficiently varied to match the combinatorial possibilities of every distinct aspect of one's primary experience.⁸⁰

The problem here is the almost total reliance on self-referential data structures of one sort or another as the generating instrument of phenomenal self-awareness. For not only is it the case that intentional data structures cannot cause first-order content to become conscious simply as a consequence of their being directed toward that first-order content (the problem of the rock), but neither can they reference themselves in an actively reflexive manner. This has to do with the direction of intentionality, the direction of cognitive reference in any and every representational data structure is antireflexive, referring away from the representing structure towards that which is represented, towards the content the state is about. It is this very other-directedness or "aboutness" which makes transparency (ignorance of the fact there is a representing vehicle) possible. Intentional data structures do not and cannot reference themselves in a directly reflexive manner because the direction of intentionality does not allow it. In this regard, Kriegel writes:

⁸⁰ Byrne (1997): 116-117; Chalmers (1996): 239; Gennaro (2008): 17-18; T. Metzinger, T. (2003). *Being No One: The self-model theory of subjectivity*. Cambridge MA: MIT Press, 2003) 59-60; Diana Raffman, "On the Persistence of Phenomenology," *Conscious Experience*, ed. T. Metzinger (Schoningh: Imprint Academic 1995) 293-308; Brent Silby, "Dretske's Alternative to HOT Theories of Consciousness" (1998). URL= <http://certaindoubt.org/wp-content/uploads/2010/04/alternative-to-hot-theories-of-consciousness-by-brent-silby.pdf>

[F]or a content to be conscious is not at all for it to be represented by itself, or for the mental state that carries it to represent itself to carry it, or anything in the vicinity . . . So it is simply false that a mental state's representation of itself is conscious in virtue of the state representing itself to represent itself.⁸¹

It may be reasonably concluded, then, that higher-order constructs do not arise to match first-order content one-on-one, and even if they did, it would lack the requisite capacity to give rise to consciousness. Higher-order theory is incapable of explaining conscious reflexivity.

3. Intrinsic or Self-Representational Theories of Reflexivity

Higher-order theory fails not because it has misconceived the primary reflexive characteristic of consciousness, but essentially because of its reliance on a two-state relational structure to achieve said reflexive self-awareness: where neither primary nor higher-order state is itself conscious, there simply is no possibility of generating the required reflexive awareness. What is needed, clearly, is not a representing of other things by the higher-order construct, not a subject-object, two-state structure, but a single-state, subject-subject structure, a representing of a cognitive state by itself to itself, a genuinely self-representational reflexive state. "Intrinsic" or "self-representational" (SR) theories insist that consciousness is indeed a matter of self-awareness not other-awareness; that

⁸¹ Kriegel (2009):126-7, (2005): 39, 43-44. 48, 50; and cf. Zahavi (1999): 6ff; Joseph Levine, *Purple Haze* (New York: Oxford University Press, 2001) 171-3.

consciousness is not a two-state cognitive construct, but a single reflexive self-representational state. The locus classicus for this position is usually traced to

Franz Brentano:

[Every conscious act] includes within it a consciousness of itself. Therefore, every [conscious] act, no matter how simple, has a double object, a primary and a secondary object. The simplest act, for example the act of hearing, has as its primary object the sound, and for its secondary object, itself, the mental phenomenon in which the sound is heard.⁸²

Moreover, he insists, the primary act of hearing and secondary awareness of that act inhere within a single mental state:

The presentation of the sound and the presentation of the presentation of the sound form a single mental phenomenon.... In the same mental phenomenon in which the sound is present to our minds we simultaneously apprehend the mental phenomenon itself. What is more, we apprehend it in accordance with its dual nature insofar as it has the sound as content within it, and insofar as it has itself as content at the same time.⁸³

This Brentanian notion of a single state with two parts, one directed outwards toward the world, and other inwards towards the state itself, has been recently adopted by several analysts.⁸⁴

⁸² Franz. Brentano, *Psychology From an Empirical Standpoint*, Eds O. Kraus, L.L. McAlister. Trans. A.C. Rancurello, D.B. Terrell, & L.L. McAlister (London: Routledge and Kegan Paul, 1874/1973) 153-4; F. Brentano, *Descriptive psychology*, B. Müller (Ed, trans) (London: Routledge, 1882) 25-26.

⁸³ Brentano (1874/1973): 127.

⁸⁴ Janzen (2008); Kriegel (2007a), (2009); Natsoulas (1996-2006, *see below*); Smith (2004); Thomasson (2000, *see below*); Van Gulick (2004), 2006. Thomas Natsoulas has authored a series of 13 articles entitled "The Case for Intrinsic Theory" (1-13), published 1996-2006. For the full list see

The self-representational approach to explaining conscious reflexivity is not without its problems, however. First and foremost, is the fact that intrinsic theories are qualia theories which explain consciousness in terms of a reflexive-awareness-of-content structure, rather than reflexive-awareness as such.⁸⁵ Van Gulick and Gennaro both explicitly criticise any deviation towards a notion of self-awareness as distinct from qualia.⁸⁶ But as the evidence reviewed above in relation to HOR theories indicates, consciousness is not qualia, reflexive self-awareness need not and does not arise as a result of the generation of perceptual, emotional or conceptual content. Moreover, to the extent that consciousness supervenes on the representation of state rather than content properties (see sec. 3 below for further discussion), the presence or absence of first-order representational content is irrelevant to and does not impinge upon the overall cognitive quality (conscious or nonconscious) of the processing state.

In the case of SR theory, with its emphasis on a single, internally complex state construct, the problematic result of conflating consciousness with

(<http://philpapers.org/s/Thomas%20Natsoulas>) in which he discusses various facets of intrinsic, “self-representational” theory, but proposes no specific psychological mechanism for operationalizing this self-representational capacity. Thomasson (2000), though nominally following Brentano, advocates a radical reinterpretation wherein the crucial self-awareness component of a conscious mental act is dropped and replaced with an adverbial notion (experiencing consciously). The result is a “conscious” experience “not in the sense that I am aware of it” (2000:203), which, lacking awareness, contravenes the conventional understanding of the term consciousness, and Brentano’s understanding in particular (as discussed above).

⁸⁵ Brentano (1874/1973): 139ff ; Janzen (2008): ch.6; Kriegel (2003b): 125, (2005): 46, 51, (2007a): 901, (2009):105-106; Smith (2004):84, 98, 100; Van Gulick (2004): 73, (2006):14; Natsoulas (1996):268, T. Natsoulas, “The case for Intrinsic Theory 12.: Inner Awareness Conceived as a Modal Character of Conscious Experience,” *Journal of Mind & Behavior* XXVII (2006): 183-214, 167.

⁸⁶ Van Gulick (2004):.73; and cf. Gennaro (2008), (2006): 226ff .

qualia is that it leads to a fundamental self-contradiction; for while all current SR theories explicitly acknowledge both nonconscious as well as conscious modes of cognitive processing,⁸⁷ the intrinsic self-representational position is essentially committed to the notion that all cognition is inevitably conscious. Intrinsic theory insists that there is a single state wherein the being-aware-of-itself is intrinsic to and thus an ineluctable consequence of the first-order content state. Conscious self-awareness arises simultaneous with and as part of the first-order cognitive act, so that the resulting state is described as being aware of its object and of itself all at the same time and as of the same, singular, state. The first-order mental occurrence and second-order awareness of it, are, as Natsoulas puts it, “a single unmediated unity [where] neither of them has any existence apart from the other.”⁸⁸ But then, if self-awareness is so utterly intrinsic to and co-existent with first-order content, how is it possible for first-order states to arise non-consciously, without self-awareness, as they clearly do most of the time.⁸⁹

⁸⁷ Janzen (2008): 21ff ; Kriegel (2003a): 1; Natsoulas (2006):185; Smith (1988): 26, (2004): 82; Van Gulick (2004): 69-70, 74, (2006):, 2; U. Kriegel,). “Consciousness as Sensory Quality and as Self-Awareness,” *Phenomenology and the Cognitive Sciences II* (2003a):1-26.

⁸⁸ Natsoulas (1996): 277; cf. Janzen (2008): 23; Kriegel (2009): 15, 114, 146; (2003b): 112; Smith (2004): 99; Van Gulick (2004): 84.

⁸⁹ A point made by P. Carruthers, *The Architecture of the Mind* (New York: Oxford, 2006) 299; cf. Rosenthal (1997): 745. The implication that all cognition is conscious may be an unfortunate legacy of adopting a Brentanian model of self-representation, since Brentano himself was convinced that all cognition is conscious: “every mental act . . . includes within it a consciousness of itself” (1874/1973, p. 153). But the evidence does not support Brentano. The cognitive unconscious is a fact of mental life, and SR theory does not allow for this fact. Nor is it possible to argue that self-awareness does not necessarily arise with every first-order cognitive event, but when it does arise it does so simultaneously with and as part of that first-order event. If self-awareness does not arise in some cases, then some additional causative factor must be at work in those cases where it does, and that additional factor is unaccounted for in self-representational theory. If it were introduced, it would be extrinsic to the first-order state which

In addition to its inability to allow for nonconscious cognition, SR theory is also vulnerable in relation to the adequacy of its accounting for consciousness itself. What concrete mechanism is being proposed to account for the reflexivity of the conscious component of the complex cognitive state? How is this self-representational capacity achieved? Janzen (2008) and Natsoulas (1996-2006) don't nominate a specific psychological mechanism. Van Gulick, on the other hand, enlists "enhanced subjectivity" (enhanced first-person or egocentric perspective) as the psychological means of achieving reflexive cognition. He writes, "[T]he reflexive meta-intentionality associated with conscious states . . . derives not from the addition of distinct explicit self-ascriptive meta-states, but from the implicit self-perspectuality [subjectivity] that is built into the intentional structure of conscious experience itself."⁹⁰ But as we have seen, subjectivity, while certainly a necessary part of reflexive consciousness does not, of itself, entail reflexivity because all cognitive acts—conscious and unconscious—are realized subjectively.

Woodruff Smith⁹¹ offers an enhanced version of Husserl's temporal layering of experience which consists of a synchronic perceptual moment, linked with a secondary retention of the just-past moment as well as anticipation

ruins the intrinsic-ness of the theory. The dilemma is somewhat akin to the infinite regress problem attached to HOR theories where the higher-order construct must remain nonconscious, for if it were claimed to be conscious, it would require a yet higher construct to make it so, and so on ad infinitum.

⁹⁰ Van Gulick (2004): 84.

⁹¹ Smith (2004): 104.

(“protention”) of the next moment.⁹² Smith retains this temporal layering structure (current *intentional* moment, retention of just past moments and protentions of future moments) but he notes⁹³ that the proper character of the conscious modality (inner awareness) is reflexivity, not mere temporal extension. Consequently, Smith proposes⁹⁴ that consciousness supervenes not on the temporal flow of representational content (past, present & future musical tones) but on the IRP temporal extension of the representational vehicle (past, present and future *hearings of* the musical tones). Now it is certainly true that the temporal layering of intentional representations to produce a diachronically unified temporal flow of experience (a flow of musical tones) does not reverse the direction of intentionality at any individual moment to generate the requisite reflexivity, as Smith notes.⁹⁵ But it is difficult to discern how reflexive awareness would arise merely on the basis of the temporal extension of

⁹² As laid out by Edmund Husserl, *On the Phenomenology of the Consciousness of Internal Time*. Translated by John Barnett Brough (Dordrecht: Kluwer, (1893-1917/1991). On which see D. Zahavi, *Husserl's Phenomenology* (Stanford CA: Stanford University Press, 2003). Smith is not alone in adopting a Husserlian explication of the mechanism of conscious awareness. Zahavi also extends Husserl's explanation of time consciousness (protention-intention-retention) to an explanation of the mechanism of consciousness as such (D. Zahavi, “Inner (time-)consciousness,” *On time—New Contributions to the Husserlian Phenomenology of Time*, eds. D. Lohmar & I. Yamaguchi (Dordrecht: Springer, 2010) 319-340. Natika Newton and Ralph Ellis contend that the ‘temporal thickness’ of conscious awareness derives from a weaving together of new sensory input with the memory of immediate past input, along with anticipations (expectations) of immediate future input. (N. Newton, “Consciousness, Qualia and Re-entrant Signalling,” *Behavior and Philosophy* XIX (1991): 21-41; cf. R.D. Ellis & N. Newton, “Three Paradoxes of Phenomenal Consciousness: Bridging the Explanatory Gap,” *Journal of Consciousness Studies* V (1998): 419-442. 425). Paula Droege also defines her conscious “now” moment as a combination of the immediate past, anticipated future and present moments (Droege, 2003: 84-88). Dan Lloyd also follows the Husserlian notion of consciousness as temporally-extended cognition. D. Lloyd, *Radiant cool: A Novel Theory of Consciousness* (Cambridge, MA: MIT Press, 2004) 272.

⁹³ Smith (2004): 104.

⁹⁴ *ibid.*, 105-6.

⁹⁵ *ibid.*, 104-105.

nonreflexive experience, nonconscious cognitive moments. The problem of the rock applies with equal force here as it did for the HO theory—two, or in this case three, nonconscious constructs don't make a conscious moment simply because they succeed or overlap each other (the problem of the *diachronic* rock).

The other prominent SR theorist, Uriah Kriegel, has developed a version of self-representation involving “cross-order integration” (COI) whereby a first-order object representation is combined with a second-order representation of that first-order state. He describes this internally complex one-state construct as follows:

[A] state is conscious because it has the right sort of representational content: it folds within it [A] a representation of an external object and [B] a representation of that representation.⁹⁶

Kriegel's contention is that the *integration* (and not simply the simultaneous activation) of parts A and B, creates an entirely new state-of-integration C, which constitutes the conscious state because it involves a state totally aware of itself in virtue of indirectly (implicitly) representing the whole of itself.

[C]onscious states have a part [B] that represents directly another part [A], and in virtue of that represents indirectly the whole state [C].⁹⁷

⁹⁶ Kriegel (2005): 46; cf. Kriegel (2009): 233, Kriegel, “The Same Order Monitoring Theory of Consciousness: Second Version,” *Synthesis Philosophica* XLIV (2007b): 361-384, 364-5

⁹⁷ Kriegel (2009): 215; cf. Kriegel (2009):.204, (2007a):.4; (2005):.48.

This emphasis on the derivation of consciousness through an indirect representing of the whole (the first-order element, the higher-order representation of that first order element, and the integration of the two) in virtue of its representing a part (the first-order state) constitutes an extension of the original cross-order integration thesis which claimed that consciousness derives from the binding of a first order sensory quality (representing an environmental feature) and a second-order representation of that first order state,⁹⁸ and, it should be noted, a complete reversal of his original insistence (2003a) that sensory quality is neither sufficient or even necessary for inner awareness.⁹⁹

In any event, the central question remains as to whether this final indirect self-representation configuration of Kriegel's COI theory provides a viable account of reflexivity. Critics have been unanimous in their assessment that it

⁹⁸ Kriegel (2003b): 1 25, (2005): 46, 51, (2007a) 901; and cf. Weisberg (2008): 166-7.

⁹⁹ As Gennaro notes (2006:235-236, 2012: 3), Kriegel's support for the equation of consciousness with qualia (his cross-order integration construct involving the integration of awareness with first-order data content) stands in flagrant contradiction to his 2003a article devoted entirely to the proposition that consciousness is not qualia, not reflexive awareness of first order content ("sensory quality"), but inner self-awareness as such. His conclusion to that 2003 article reads: "[A]ccording to our account, for one's experience to be conscious at all, it must involve implicit self-awareness; but which specific kind of experience it is, is a matter of which sensory quale one is (implicitly) aware of one's experience as having. . . The upshot is that consciousness is not a matter of sensory qualia. . . . To devise a theory of consciousness, what we would have to explain is the phenomenon of permanent implicit self-awareness" (2003a:21). Having argued at length and clearly demonstrated that sensory quality is neither sufficient nor even necessary for consciousness, his subsequent espousal of a qualia theory is inexplicable. There is, in addition, the perplexing fact that Kriegel, in his 2009 book (68-71, ff.) attempts to uphold the principle of transparency, which insists that state characteristics of an experience, such as conscious awareness itself, are "diaphanous" and unnoticed, not phenomenally explicit (you look through the state characteristics and see only the perceptual-conceptual-emotional content) while at the same time basing his theory of subjective awareness on the claim that we do, Brentano-style, experience the state of representing bluishness alongside the bluish sensory content. (2009: 109-110). The latter is Smith's claim that consciousness has to do with state (modality) characteristics and not sensory qualities (mode), which Kriegel has consistently rejected (2003b, 2005, 2009:103).

does not. They point out that (1), if Kriegel's cognitive construct did generate conscious self-awareness, it would, in fact, produce two conscious states, not the single state he claims. In chapter 3 his first-order construct "qualitative character" consists of the representation of response-dependent properties of an object. But this is emended in Ch. 4 such that "qualitative character" derives not from first-order representational content of, for example, a red tomato, but from higher-order representing of the first-order state of perceiving a red tomato.¹⁰⁰ This, in effect leads Kriegel to posit 2 conscious states;¹⁰¹ the first consisting of the aforementioned higher-order representation of the first-order state;¹⁰² and the second, of the self-awareness arising from the whole integrated state.¹⁰³ Kriegel concedes this by granting the former "submaximal consciousness" status and the latter "maximal".¹⁰⁴ These twin conscious states would, if actual, undermine Kriegel's claim to be describing a single, self-representing state.¹⁰⁵

This eventuality is somewhat mitigated by the second arm of critical assessment to the effect that (2) Kriegel's construct could not be said to give rise to phenomenally-explicit conscious self-awareness at all. Kriegel's account

¹⁰⁰ J. Levine, "Review of *Subjective Consciousness* by Uriah Kriegel," *Notre Dame Philosophical Reviews* 2010, III.

¹⁰¹ R. Lurz, "Review of *Subjective Consciousness: A Self-representational Theory*, by Uriah Kriegel," *Grazer Philosophische Studien* LXXXV (2012):.347-353, 352.

¹⁰² Kriegel (2009): 110-111.

¹⁰³ Kriegel (2009): 215-232.

¹⁰⁴ Kriegel (2009): 229.

¹⁰⁵ Kriegel (2009): 15, 114, 146; (2003b): 112.

of subjective self-awareness rests entirely on the notion of representational content. His publications have consistently rejected Woodruff Smith's position that consciousness derives from properties of the cognitive vehicle, the representational process,¹⁰⁶ and have insisted, to the contrary, that self-awareness derives from properties of the content represented.¹⁰⁷ The problem, as Kriegel points out,¹⁰⁸ is that intentional representation is always inherently antireflexive, the direction of intentionality is outward not inward, away from itself to something else, not towards itself. Kriegel frequently acknowledges this antireflexivity of intentional content in terms of the antireflexive direction of causality (a thing cannot cause itself).¹⁰⁹ He highlights the insufficiency of representing, of itself, to achieve self-awareness when he writes

[F]or a content to be conscious is not at all for it to be represented by itself, or for the mental state that carries it to represent itself to carry it, or anything in the vicinity . . . So it is simply false that a mental state's representation of itself is conscious in virtue of the state representing itself to represent itself.¹¹⁰

It is in order to circumvent the nonviability of direct self-representation that Kriegel proposes his theory of indirect self-representation (outlined above). But if direct representational content is inherently incapable of producing a reflexive cognitive gesture, as Kriegel concedes, then an indirect, merely implicit

¹⁰⁶ Smith (1986):150, (2004):84, 98ff ; D.W. Smith, "The Structure of (Self-)Consciousness," *Topoi* V (1986): 149-156.

¹⁰⁷ Kriegel (2005): 40ff, (2003a): 19, (2009): 103.

¹⁰⁸ Kriegel (2005): 31, 43-4.

¹⁰⁹ Kriegel (2005): 43-45, (2007b): 364, (2003c):. 482, 493, (2009): 205ff.

¹¹⁰ Kriegel (2009): 126-7; cf. (2005): 48, 50.

representation of the whole state cannot possibly be said to constitute a “phenomenologically manifest” state of self-awareness. This is the conclusion of analysts,¹¹¹ and Kriegel himself concedes¹¹² that indirect (implicit) representation does not become consciously explicit.¹¹³ Indirect self-representation, then, does not constitute a viable mechanism for the generation of reflexive self-awareness.

Reflexivity Derives From a Reflexive Processing Regime (not Self-Referential Data Structures)

It is becoming increasingly apparent that current attempts to explain the cognitive mechanism of reflexivity, both higher-order representation theories and intrinsic or self-representational theories, may well have reached an impasse, for lack of a viable mechanism, which is the critical enabling factor as

¹¹¹ Levine (2010) 3; Lurz (2012):352; M. Sebastian, “Subjective Consciousness: a Self-Representational Theory by Uriah Kriegel,” *Disputatio IV* (2012): 413-417, 415; R. Van Gulick, “Subjective Consciousness and Self-Representation,” *Philosophical Studies CLIX* (2011): 457-465, 464; J. Weisberg, “Review of *Subjective Consciousness: A Self-Representational Theory*, by Uriah Kriegel,” *Mind CXX* (2011): 538-542, 541.

¹¹² Kriegel (2009): 230.

¹¹³ Kriegel’s admission (2009:230) that indirect (implicit) representation of the whole cognitive state is not phenomenally explicit (i.e. not conscious) is similar to an earlier admission (2005:50) that his cross-order integration construct does not achieve self-representational consciousness. He wrote: “the cross-order information integration approach offers, in fact, a representationalist account of subjective character but one which does not appeal to self-representation. Instead, it suggests a way in which mental states that are not self-representing could have the sort of representational content that self-representing states would have. . . . It would be overly presumptuous to conclude from these remarks that cross-order information integration *must* be the key to the subjective character of conscious experience.”

Gennaro rightly suggests.¹¹⁴ And this for two reasons, both of which derive from the almost universal reliance on intentional data structures of one sort or another as opposed to the way data structures are processed. The first problem is that for intentional data structures, the direction of intentionality, the direction of cognitive reference in any and every representational data structure is antireflexive. Consequently, even elaborate representational contortions (Husserlian iterations, Kriegel's implicit re-representations) do not succeed in overcoming this inherent antireflexivity of intentional representation. The second problem for current theories of reflexivity is that while intentional data structures cannot reference themselves in an actively reflexive manner, neither can they cause first-order content to become conscious simply as a consequence of their representing or being directed toward that first-order content.¹¹⁵

But if intentional data structures can neither reference themselves reflexively nor confer reflexivity on another, first-order structure, how is reflexivity to be achieved? The solution may lie not in the data structure but in the way it is processed. Both Gennaro¹¹⁶ and Kriegel emphasize the capacity of a cognitive processing regime to create a genuinely new cognitive event. Kriegel writes "if two states are united by a psychologically real process, that

¹¹⁴ Gennaro (2006):226ff.

¹¹⁵ Known as "the problem of the rock," discussed above *Higher-Order Theories of Reflexivity*; and see Levine (2006): 22-23.

¹¹⁶ Gennaro (2006): 237-240.

process would presumably affect the causal powers of the whole.’’¹¹⁷ The particular mechanism they have both proposed, the integration of two data structures, is not actually associated with conscious processing—masked priming and subliminal perceptual integration both remain unconscious—and therefore integration is insufficient to the task.¹¹⁸ Furthermore, even if integration were equal to the task, it is unclear how much of the phenomenological workload (generating self-awareness) is being performed by the processing regime and what percentage devolves upon the fact that the secondary data structure represents the primary act of world representation.¹¹⁹

Clearly, not just any processing regime will generate the requisite reflexive state. But an immediately recursive processing regime which essentially monitors or cognizes itself, most certainly could underwrite a reflexive form of cognitive awareness, particularly if that recursive monitoring did not rely on a reading of content properties but rather on the registration of state properties, the chief feature of which in this case is the reflexive auto-noetic quality of the processing regime. There is good evidence for both capacities in the cognitive system.

¹¹⁷ Kriegel (2009): 224, n.37; cf. (2005): 49-51, (2007b): 901; Schlicht (2011).

¹¹⁸ Weisberg (2008); 173.

¹¹⁹ This is Smith’s argument (2004:104-5) for the insufficiency of Husserl’s temporal fusion of retentions, intentions and protensions into an integrated, singular, flowing state to produce reflexive conscious awareness. It is arguable, though, that integration plays such a significant role in Husserl’s construct as to include it with the theories of Gennaro and Kriegel as a (somewhat) process-oriented theory.

As to the first, the viability of reflexive processing, we have discussed above the self-referential character of cognitive architecture (serving the basic requirement for self-regulatory engagement with the environment), and the fact that cognitive systems have developed an even more proactive feed-forward or anticipatory form of self-reference in the form of predictive emulation architecture. This development is significant because predictive self-referential processing provides the basis for developing the capacity for the self-referential monitoring of a process by itself simply by rationalizing existing loops that update a current state with a predicted next state, into a more immediately recursive loop that updates a current state by a virtual copy of itself (see Peters, 2010, for detailed discussion). Predictive architectures already employ a more temporally extended form of recursion (recurrent self-reference) as a way of monitoring the capacity of motor outputs to achieve desired perceptual inputs. Rationalizing this periodically self-referencing anticipative circuitry into a more immediately recursive, self-updating circuit simply repeats the original evolutionary emergence of fast predictive processing loops within slower motor-output-to-perceptual-feedback loops that form the basis of predictive processing architecture.

Recursive self-activation (or self-updating) at the neural level has the capacity to support reflexive self-knowing or self-awareness at the cognitive level, on the basis that reflexive self-awareness embodies a registration of state rather than content properties; in this case the reflexivity of the processing

regime. The capacity of the cognitive system to register features of its processing state as well of the content of that state has received a great deal of attention, and specifically in relation to the question of consciousness. Several analysts point out that our conscious experiences do explicitly register qualitative features that are not identical to the particulars of the objects represented.¹²⁰ These include the “inner light show” one experiences when one presses a finger against the eyeball,¹²¹ the continuous explicit awareness of the distinction between current auditory and visual streams of sensation,¹²² as well as non-representational qualities of these sensations, such as the difference between seeing clearly and with blurred vision—where blurriness is a property of the visual process not the content.¹²³ In addition, there is the direct awareness of a distinction between memories recalled to mind as against ongoing perception.¹²⁴

¹²⁰ Janzen (2008):142; Smith (2004):99; Zahavi (1999):23 ; R. Gennaro, “Representationalism, Peripheral Awareness, and the Transparency of Experience,” *Philosophical Studies* CXXXIX (2008b): 39-56, 47.

¹²¹ Amy Kind, “What’s so Transparent About Transparency?” *Philosophical Studies* CXV (2003): 225-244, 237

¹²² Michael Pace, “Blurred Vision and the Transparency of Experience,” *Pacific Philosophical Quarterly* LXXXVIII (2007). 328-354, 332; A.L. Thomasson, “Phenomenal Consciousness and the Phenomenal World,” *Monist* XCI (2008): 191-214, 205.

¹²³ Janzen (2006):326; Kind (2003):12; Pace (2007); Paul Boghossian & J.David Velleman, “Colour as a Secondary Quality,” *Mind* LXXXVIII (1989):81-103.

¹²⁴ Often styled “reality monitoring” or “source monitoring.” See Kunzendorf (1987-88); Marcia Johnson, “Reality Monitoring: Evidence from Confabulation in Organic Brain Disease Patients,” *Awareness of Deficit After Brain Injury. Clinical and Theoretical Issues*, eds. George Prigatano & Daniel Schacter (New York: Oxford University Press, 1991) 176-197. Marcia Johnson, Shahin Hashtroudi & D.Stephen Lindsay, “Source Monitoring,” *Psychological Bulletin* CXIV (1993): 3–28.

The temporal dimension also expresses the registration of state rather than content properties. During the passage of the sensory, emotional and conceptual events, there is ongoing, overt awareness of the temporal duration of an experience, the passage of time, a temporal awareness which is intrinsic to the cognitive state, not the objects represented in that state. This temporal awareness is called subjective time because time is not a quality directly registered by the senses, but constructed internally. Of significance is the fact that this internally-constructed sense of duration varies. Time spent in interesting and novel surroundings that one is attending to and actively exploring can seem like a very long time. Acutely life-threatening situations can slow time seemingly to a standstill. The course of an average undemanding, uneventful day, on the other hand, can flow by relatively quickly. This difference in the sense of time passing quickly or slowly is related, as Pockett (2003) explains, to a difference in the duration of *now*. Experimental studies suggest that the subjectively experienced duration of *now* can vary from milliseconds,¹²⁵ through hundreds of milliseconds,¹²⁶ to one or two seconds.¹²⁷ The duration of this now-moment, in turn, is a direct reflection of the rate of sensorimotor sampling of the external world, or better, according to the rate of

¹²⁵ Sue Pockett, "How Long is 'Now'?" *Phenomenology of the Specious Present*, *Phenomenology and the Cognitive Sciences II* (2003):55-68.

¹²⁶ Tallis. Bachmann, *Microgenetic Approach to the Conscious Mind*. Amsterdam: John Benjamins, 2000).

¹²⁷ Ernst Poppel, & Tom Artin, *Mindworks: Time and Conscious Experience* (Boston: Harcourt Brace Jovanovich, 1988).

sensorimotor processing which includes sampling (Pockett, 2003). The subjective sense of the duration of *now* expands and contracts as the rate of sensorimotor updating expands and contracts, but inversely; that is, a faster rate of updating generates more *now* moments in relation to the actual passage of the event - more subjective time is packed into the event - which makes it seem to be passing more slowly. Fewer updates of subjective *now* pack in less *now* moments, less time into an event which seems to pass more quickly.

The principal implication is that this sense of temporal duration reflects a registration of properties of the cognitive state (the rate of sampling which generates the state), not features of the particular objects which comprise the content of the represented event. Though it may not seem so, subjectively sensed time is actually a feature of the representing vehicle or state, not a quality or feature of the event represented much less the objects represented. Taken in conjunction with the evidence, referenced above, that the conscious state does not consist in the awareness of representational content, that conscious is not qualia, we are left with the conclusion that consciousness must reflect a reading or registration of a state property,¹²⁸ in this instance the reflexive or auto-noetic character of the state engineered by recursive processing circuitry.

¹²⁸ Smith (1986):150, (2004):84, 98ff; Janzen (2008):158-61.

Conclusion

The theory outlined above can be distinguished from existing higher-order and self-representational theories in terms of the allocation of phenomenal labour.

Where previous theories of consciousness as reflexivity rely wholly or in large measure on the self-referential intentionality of data structures¹²⁹, the theory advanced here relies exclusively on the self-cognizing capacity of a recursive processing regime, that is, on the recursive manner of processing along with the fact that cognition is reading the features of that manner of processing. This could be understood as a return to the focus of earlier psychological theories which attempted to explicate consciousness as issuing from a particular kind of processing activity, as proposed by Johnson-Laird: “if consciousness depends on the computations of the nervous system, then it is likely to be a property of the algorithms that are used to carry out those computations rather than a property of their [informational] results.”¹³⁰ Suggested processing mechanisms include the coordination of a central executive, attentional highlighting of informational content, internal linguistic commentary on perception, the comparison of sensory input with stored memory engrams, short term or working memory, the global broadcasting of information, the logical computation of meaningful symbols, action selection, source monitoring, and

¹²⁹ Rosenthal’s higher-order thought theory (1997), for example, relies totally on the referential content of data structures to generate the requisite “awareness that,” while Gennaro (2006) and Kriegel (2007a) introduce some measure of reliance on the processing regime involved.

¹³⁰ Phillip Johnson-Laird, “A Computational Analysis of Consciousness,” *Consciousness in Contemporary Science*, eds. A.J. Marcel & E. Bisiach (New York: Oxford University Press, 1988) 357-368, 360.

recurrent (but not immediately recursive) processing loops.¹³¹ Interestingly, cognitive psychologists and neuroscientists¹³² have also proposed mechanisms involving the embedding of a first-order data structure within a secondary representation of that first order process or state, a construct which is similar to those of Kriegel and Gennaro. Flohr envisages the process as follows:

The system would generate second-order representations of internal states. An iteration of such processes would generate higher-order, self-referential representations of the system itself and of its current state [of representing]. The system can . . . embed first-order representations in a model of itself and thereby

¹³¹ On central executive coordination as consciousness, see Carlo Umiltà, “The Control Operations of Consciousness.” In Marcel & Bisiach (1988) 334-356. For Consciousness understood as attentional highlighting of informational content, see Neisser, (2006); Velmans (1991); and Eric Harth, “*The Creative Loop: How the Brain Makes a Mind* (New York: Addison-Wesley, 1993); Consciousness is styled as internal linguistic commentary on perception by Otto Creutzfeldt, “Neurophysiological Mechanisms and Consciousness, *Brain and Mind. Ciba Symposium 69* (New York: Elsevier, 1979); Julian Jaynes, *The Origin of Consciousness in the Breakdown of the Bicameral Mind* (London: Penguin, 1993); Alexander Luria, “The Human Brain and Conscious Activity,” *Consciousness and Self-Regulation*, Vol. 2, eds. Gary Schwartz & David Shapiro (New York: Plenum, 1978) 1-35); David Olds, “Consciousness: a Brain-Centered, Informational Approach,” *Psychoanalytic Inquiry* XII (1992):419-444. Consciousness is cast as the comparison of sensory input with stored memory engrams by Jeffrey Gray, “The Contents of Consciousness: a Neuro-Psychological Conjecture,” *Behavioral and Brain Sciences* XVIII (1995):659-722.); as short term or working memory in Johnson-Laird (1988), and Richard C. Atkinson & Robert M. Shiffrin, “The Control of Short-Term Memory,” *Scientific American* CCXXIV (1971): 82-90; as the global broadcasting of information in Baars (1988); as the logical computation of meaningful symbols in Jerry Fodor, *The Language of Thought* (Cambridge MA: Harvard University Press, 1975.); as action selection by Tim Shallice, “Dual Functions of Consciousness,” *Psychological Review* LXXIX (1972): 383-393; as source monitoring in Kundendorf (1987-88); and finally, as recurrent (but not immediately recursive) processing loops in Victor Lamme, “Towards a True Neural Stance on Consciousness,” *Trends in Cognitive Sciences* X (2006):494-501; Rodney Cotterill, “Prediction and Internal Feedback in Conscious Perception,” *Journal of Consciousness Studies* III (1996):245-266; Gerald Edelman, *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989). In all of these psychological models, it is the processing regime that generates conscious awareness.

¹³² Cognitive psychologists include Johnson-Laird (1988) and Hans Flohr, “Sensations and Brain Processes,” *Behavioural Brain Research* LXXI (1995):157-61.). Embedding has also been proposed by neuroscientist Antonio. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace & Company, 1999).

represent itself as an actually representing system . . . they would know ‘what it is like’ to be such a system.¹³³

With this in mind, the trend, in the theories of Gennaro and Kriegel, toward a greater emphasis on process over intentional representation expresses a move by philosophy of mind toward a more psychological and empirically verifiable approach to the explanation of consciousness. Certainly this is the intention of the theory outlined here, where a fully recursive processing regime generates a reflexive state.

In the pantheon of consciousness theories, this solution would best be described as a *reflexive state* theory. As opposed to theories which envisage consciousness as qualia, wherein consciousness arises either as intentional representation per se (FOR theories), or as intentional representation that is object of higher order representation (HOR theories), or a self-representing element of a complex state (Intrinsic Self Representational theories), reflexive state theory characterizes consciousness as an actively self-cognizing state *extrinsic* to primary informational streams. Rather than the self-representational cognitive structure being conjoined to, arising as part of, or being directed at primary informational content in some fashion, the direction of fit is reversed: conscious self-awareness is independently generated and that portion of unconscious processing which is directed to (or permitted access into) this state

¹³³ Flohr (1995):160.

becomes incorporated¹³⁴ into the reflexive state and consequently partakes of the reflexive quality of the state. But regardless of the amount or type of primary information currently incorporated, the reflexive state remains a cognitive entity distinct from that content because independently constituted as a consequence of the registration or reading of state rather than of content properties. Consequently, the principal property of a recursive processing state is reflexive or auto-noetic awareness—consciousness.

¹³⁴ Other incorporational models include those of Metzinger (2000), Dennett (1991), Van Gulick (2004), and Antti Revonsuo, *Inner Presence: Consciousness as a Biological Phenomenon* (Cambridge MA:MIT Press, 2005). But all envisage consciousness as something other than a reflexive auto-noetic state. The first two envisage information becoming conscious when incorporated into a spatiotemporal array which is somehow “conscious” in virtue of its spatiality (Revonsuo (2005):123) or its status as the highest order model (Metzinger (2000):. 289-290, 299). Daniel Dennett’s 1991 “multiple drafts” theory of consciousness models the recruitment of content into an ongoing serial narrative, which for Dennett, as for Metzinger and Revonsuo, does not amount to a reflexively self-aware state, but is nominated “conscious” nonetheless. For van Gulick, individual phenomenal elements are subsumed within the larger conscious state, which, however, is conscious simply in virtue of being subjective (Van Gulick (2004):19-20, (2006):30). However, as discussed above (*Reflexivity as Subjectivity*), subjectivity characterizes all cognition but is insufficient to explain reflexive conscious cognition.