



1 *tion, and for the way in which we are aware of changes, transitions,*  
2 *and boundaries between mental states in specific cases of intro-*  
3 *spective consciousness.*

4 **Keywords:** higher-order representation; introspective consciousness;  
5 stimuli-induced; self-triggered; Higher-Order Representation Inten-  
6 tionally for Introspective Consciousness.

## 7 **1. Introduction and** 8 **Preliminary Considerations**

9 Let us suppose that you are feeling a severe pain in your toe right  
10 now. In fact, I am stepping on it while we are standing together.  
11 Besides no doubt being in a state of pain, you also feel confused and  
12 upset since I am roaring with laughter as I stomp on you. The pain  
13 sensation in your toe and the feelings of confusion and upset occupy  
14 centre stage in your stream of consciousness seemingly capturing your  
15 awareness of such mental states — as occurring to you — without any  
16 effort on your part. Take another case. Let us suppose that you have  
17 taken your glasses off and right away the text on the computer screen  
18 becomes blurry. You are not often tempted to think that the screen —  
19 i.e. the physical object — is as such blurry. It is the visual experience  
20 of the screen that is blurry, not the screen itself. In such a case you  
21 might spontaneously report: ‘I cannot see the screen’, and be aware of  
22 your visual experience.

23 We take these scenarios to be typical cases of introspective con-  
24 sciousness. In the simplest terms, introspective consciousness is a way  
25 to be aware of your own current and recently past mental states, and to  
26 self-attribute those mental states to yourself. There seem to be cases  
27 where you engage in introspection spontaneously as in the probing of  
28 your pain sensation because of my stomping on your foot, *and* cases  
29 where you do it purposefully, as when your own interest or volition  
30 initiates introspective consciousness of an occurring mental state. In  
31 the above example, if I am stomping on your foot to teach you how to  
32 introspect painful experiences, then at first you may spontaneously  
33 introspectively access the pain sensation. But afterwards you may  
34 purposefully direct your powers of introspection just to the pain sensa-  
35 tion that you are feeling right now, instead of another state such as a  
36 desire to take revenge.

37 Now, if I ask further how you know that you are having a first-order  
38 state — i.e. feeling or experiencing such a stabbing pain or seeing  
39 blurry — it is likely that you will express that you have a subjective

1 conviction that — depending on the mental state you are in — you are  
2 feeling a stabbing pain in your toe or seeing the screen as blurry.  
3 Notice that in asking the question ‘how do you know that you are  
4 feeling that pain sensation or having such a visual experience?’, I am  
5 asking for your grounds for judging that you are in that specific  
6 mental state. In other words, I am asking what justifies your holding  
7 such judgments or self-attributions.<sup>3</sup> Further enquiring like this may  
8 require reflection or deeper degrees of introspection, maybe a third-  
9 order state, which do not occur regularly or frequently — issues that  
10 we leave for another discussion.

11 What the previous cases suggest is that introspection can be of two  
12 different modes: either ‘stimuli-induced’ or ‘self-triggered’ (Renero,  
13 2019). A severe pain caused by my stomping on your foot will lead  
14 you to introspectively access your pain sensation in a seemingly  
15 involuntary way. Here, an episode of introspective consciousness is  
16 induced by the stimuli or the pain itself — we shall call this *stimuli-*  
17 *induced introspective consciousness*. In the other case, you voluntarily  
18 engage in probing your own mental state occurring. Here, an episode  
19 of introspective consciousness is initiated in a voluntary way by  
20 selecting a mental state to focus on: either your pain sensation or your  
21 emotional response to it, instead of your visual state of your swollen  
22 toe, for example — we shall call this *self-triggered introspective*  
23 *consciousness*.

24 Our goal in this work is to provide the grounds for a novel theory of  
25 introspective consciousness that can allow for both of these modes of  
26 introspective access *and* to account for the way in which we are aware  
27 of changes, transitions, and contrasts between mental states in specific  
28 cases of introspective consciousness. To do this, we will build on a  
29 version of the higher-order theory of consciousness called the Higher-  
30 Order Representation of a Representation (HOROR) theory of phen-  
31 omenal consciousness (Brown, 2015) aligned with a pluralist view on  
32 introspection (Renero, 2019; 2017). We call this novel view: *Higher-*  
33 *Order Representation Intentionally For Introspective Consciousness*  
34 (HORIFIC).

35 Although HOROR has been inspired by the well-known higher-  
36 order thought theory of consciousness — i.e. the HOT theory  
37 (Rosenthal, 1986; 1993; 1997; 2005; see also Gennaro, 1993; 1996;

---

<sup>3</sup> Whether such judgments are necessarily describable or verbally reported requires a different discussion that we cannot address here.

1 2004) — and HOROR could be considered a version of the HOT  
2 theory, we will not engage here in a detailed comparison or contrast  
3 between HOROR and HOT theories (see Berger and Brown, 2021, for  
4 a comparison) or on the benefits of higher-order theories over first-  
5 order theories (see Lau and Rosenthal, 2011; Brown, Lau and  
6 LeDoux, 2019; Brown, LeDoux and Rosenthal, 2021). We take the  
7 HOROR theory as an empirical conjecture about the nature of phen-  
8 omenal consciousness (Brown, 2014; Brown, Lau and LeDoux, 2019).  
9 Although the jury is currently out, we think that the higher-order  
10 approach to consciousness is a viable contender and should be  
11 explored and developed in enough detail that it can face the tribunal of  
12 experience.

13 The rest of the paper is divided into four main sections: §2 provides  
14 some reasons for looking beyond the traditional higher-order account  
15 of introspection; §3 introduces the basics of HOROR theory which  
16 serves as the basis to develop our view; §4 introduces our novel view,  
17 HORIFIC, and its main characteristics; §5 highlights some of the  
18 merits of HORIFIC by demonstrating this view at work and closes by  
19 offering a summary.

## 20 **2. Moving Beyond Traditional** 21 **Higher-Order Views of Introspection**

22 In this section we offer some reasons to look beyond the traditional  
23 higher-order view to account for introspection. One very general  
24 reason for this enquiry is that introspection has been relatively under-  
25 explored in the accounts offered by higher-order theories — even  
26 though introspection is considered a ‘special case of consciousness’  
27 and ‘a more complex phenomenon’ (Rosenthal, 2005, pp. 27–9).  
28 Several authors (e.g. Shargel, 2016; Berger, 2017; Carruthers and  
29 Gennaro, 2020) seem to take for granted that any higher-order theory  
30 of consciousness will adopt the model of Rosenthal (2005) and  
31 Gennaro (2012) where introspection consists in having a conscious  
32 higher-order thought. The HORIFIC view contributes by exploring  
33 and accounting for this special case of consciousness from a higher-  
34 order representation perspective.

35 Another reason, related to the first, is that philosophical work on  
36 introspection is often presented exclusively as some kind of ‘self-  
37 monitoring’ or ‘self-scanning’ process targeting a current state  
38 (Armstrong, 1968/1993; Rosenthal, 2005; Gennaro, 2012), which sets  
39 aside some kind of knowledge by ‘acquaintance’ or direct awareness

1 to one's mental states (Chalmers, 1999; Gertler, 2011; 2012).  
2 Although introspection can be considered a self-monitoring process,  
3 we find that there is something to the phenomenology of acquaint-  
4 ance. There is a sense in which it seems to me that I can have a kind of  
5 direct access to my experience or to the phenomenal character of my  
6 conscious state, and thus become aware of its nature or, at least, some  
7 of its properties. We think that this phenomenology of acquaintance  
8 can be accounted for by moving from the traditional higher-order  
9 thought view to the HOROR view. The HORIFIC view we develop  
10 can also capture the phenomenology of introspection as the content of  
11 introspective higher-order representations of first-order representa-  
12 tions. It is precisely this phenomenology of introspection that the  
13 traditional higher-order account of introspection cannot capture.  
14 According to the traditional account, having a conscious higher-order  
15 thought — i.e. having a third-order thought representing oneself as  
16 thinking a thought about one's first-order mental states (*cf.* Rosenthal,  
17 2005, p. 48; Gennaro, 2012, pp. 56–8) — will make it appear from  
18 one's point of view that one is having a conscious thought. That is to  
19 say, one will experience consciously thinking that one is seeing red.  
20 This is just the result of a kind of self-monitoring or self-scanning  
21 process. But consciously thinking a thought such as 'I am seeing red'  
22 is not the same as experiencing being acquainted with a phenomenal  
23 property.

24 In fact, according to HOT theory, introspection is having conscious  
25 thoughts that one is in particular mental states. But those thoughts do  
26 not reveal the nature of those states, nor can we infer them from the  
27 close relation between introspection and its content (see Rosenthal,  
28 2005, pp. 43–4). We think that the subject can form representations —  
29 instead of thoughts — and although we agree that inference is a  
30 different mechanism, we think that introspection can reveal something  
31 about the nature of one's target mental states. The HORIFIC view  
32 contributes by accounting for the nature of the first-order representa-  
33 tions or certain properties of the target mental state.

34 An additional reason to look beyond the traditional higher-order  
35 view to account for introspection is that the classic model of intro-  
36 spection on HOT theory takes introspection as deliberately focusing  
37 attention to one's mental states or conscious experiences (*ibid.*, see pp.  
38 48, 108–23). On this view, when one is non-introspectively conscious  
39 one is consciously aware of the world, but when one introspects one  
40 comes to shift one's attention to one's conscious experience (*ibid.*, see  
41 pp. 103–4). It may be the case that we can have this kind of deliberate

1 and focused access, and it may fit well with our proposal of a self-  
2 triggered introspective consciousness. However, it would not capture  
3 or fit well with stimuli-induced introspective consciousness where  
4 one's awareness is drawn to the experience, not to a thought about it  
5 — this is to say that it seems to one that one is aware of the experience  
6 rather than having a conscious thought (see above our second reason).

7 Furthermore, the HORIFIC view contributes by proposing an  
8 account which includes both the stimuli-induced and self-triggered  
9 modes of introspective consciousness — a distinction that has been  
10 neglected in the philosophical literature on introspection. This account  
11 builds on HOROR theory and aligns with a pluralist theory of intro-  
12 spection, it also extends them by advancing proposals specific to this  
13 special case of consciousness. The HORIFIC view extends well  
14 beyond what is present in work on HOT and HOROR theories of con-  
15 sciousness, and it could be adapted to fit with other theories of con-  
16 sciousness as well.

17  
18

### 3. The HOROR Theory of Phenomenal Consciousness

19 Since our goal is to build upon the HOROR theory of phenomenal  
20 consciousness, we will start with a very brief account of the basics of  
21 the theory. HOROR theory is a higher-order representational theory of  
22 phenomenal consciousness. *Phenomenal consciousness*, in the most  
23 general sense, is just the property a creature has of there being some-  
24 thing that 'it is like' for them (Nagel, 1974). Conscious experiences  
25 are distinguished from each other by their specific phenomenal  
26 character. In this sense, the *phenomenal character* of an experience is  
27 just the specific way that the conscious experience is like for the  
28 creature in question — e.g. consciously seeing blue, hearing a  
29 trumpet, thinking that 'cinnamon' is hard to spell, etc. This notion of  
30 phenomenal consciousness is neutral and accepted by theorists with  
31 different approaches to consciousness (e.g. Chalmers, 2018; Racciah,  
32 Block and Fox, 2021). 'Phenomenal character' captures the specific  
33 property in which conscious experiences differ from each other. A  
34 visual experience of seeing blue versus seeing green, and an auditory  
35 experience of hearing a trumpet versus hearing a double bass, will  
36 differ in their phenomenal character. For one it will be like seeing the  
37 colour quality of blue or green, and the other will be like hearing the

1 sound quality of a bright-tone trumpet or the rustling-like sound of a  
2 double bass.<sup>4</sup>

3 Providing an account of phenomenal consciousness is the primary  
4 goal of any theory of consciousness. One approach to understanding  
5 the nature of phenomenal consciousness is via *representationalism*. In  
6 its simplest form, representationalism is the view that holds that  
7 phenomenal consciousness supervenes on, or is identical with, some  
8 kind of impure representation; where that means a representational  
9 content represented in some particular way. Representationalism  
10 comes in two kinds: first-order and higher-order. On the one hand,  
11 first-order theories (e.g. Tye, 1994; Dretske, 1995) hold that phenom-  
12 enal consciousness consists in representations of properties in the  
13 environment — though see Gottlieb (2019) for an argument that these  
14 theories collapse into higher-order theories. On the other hand, higher-  
15 order theories hold that phenomenal consciousness crucially involves  
16 representations of one’s own mental life. Since we are typically aware  
17 of things either by perceiving them or by thinking that they are  
18 currently present, higher-order theories have been divided into higher-  
19 order *perception* (e.g. Armstrong, 1968/1993; Lau, 2019) and higher-  
20 order *thought* (Rosenthal, 2005) theories.

21 The HOROR theory starts with the folk platitude that consciousness  
22 involves an awareness of our mental life and then identifies the appro-  
23 priate kind of awareness with an appropriate higher-order *representa-*  
24 *tion*. Here ‘appropriate’ means that the higher-order representation  
25 subjectively appears to have arisen spontaneously, independently of  
26 any inference and that it represents oneself as currently being in the  
27 *target* mental state. It also postulates that the contents of the higher-  
28 order representations will account for the phenomenal character of  
29 one’s experience.<sup>5</sup> When one consciously sees, for example, the blue  
30 sky, one would then be aware of oneself as seeing blue. One would  
31 attribute to oneself an occurrent mental state, say, a representation of

---

4 For a view of consciousness which includes both an account of the phenomenal character of mental states and how those states become conscious — i.e. a view that builds on the HOT theory of consciousness plus the quality-space theory of mental qualities (and how these theories can work in tandem), see Renero (2014).

5 Here we use the term *target* mental state to designate the mental state which is represented by the relevant higher-order representation. We use ‘content’ of a representation to refer to the way in which the representation in question represents its target. If I am in a state which is a representation of something non-mental, then the target is the physical object and the content amounts to the satisfaction conditions placed on the targeted object.

1 seeing blue. In a nutshell, HOROR theory is a representational theory  
 2 in that it holds that phenomenal consciousness can be understood in  
 3 terms of representational states. Specifically, the HOROR theory says  
 4 that when one is phenomenally conscious, one has the appropriate  
 5 kind of higher-order representation.

6 We can put the HOROR theory of phenomenal consciousness more  
 7 specifically:

8 **HOROR:** For a subject *S* to be in a phenomenally conscious  
 9 state *C* with phenomenal character *P* is just for *S* to token an  
 10 appropriate higher-order representation with the content that *S* is  
 11 in *C*, which has character *P*.

12 As we can see, the basic idea is that, for phenomenal character, if one  
 13 represents oneself as having it, then one does in fact have it. Let us  
 14 suppose that *S* is looking at a ripe tomato right now, so *S* has a  
 15 phenomenally conscious state. The phenomenal character of *S*'s visual  
 16 state of the tomato is its redness and roundness. Notice that while a  
 17 first-order representational theory will account for this in terms of a  
 18 first-order representation of the redness and roundness of the tomato, a  
 19 higher-order representation view will account for this in terms of a  
 20 higher-order representation of oneself as being in those kinds of first-  
 21 order representational states.

22 In this case, the suitable higher-order representation will have as its  
 23 content something like the following: *I am seeing a red round object*.  
 24 That is, of course, a rough approximation of the content. But the basic  
 25 idea is that the content of the higher-order representation deploys  
 26 concepts which describe one as oneself currently *being in* the repre-  
 27 sentations which are targeted by the appropriate higher-order repre-  
 28 sentations of a representation (HORORs), and so it will seem from  
 29 one's point of view that one *is* in those mental states. The targeted  
 30 first-order representations are characterized by the targeting HOROR  
 31 as presenting objects in the environment which have properties like  
 32 colour, sound, etc. The HOROR theory is built with this common-  
 33 sense picture in mind and it is this pre-theoretic notion that the  
 34 colours, sounds, etc. seem to be out there on the objects or events that  
 35 the theory aims to account for. One does this, according to the  
 36 HOROR theory, by representing oneself as being in states that put you  
 37 in a special relation to perceptible properties. That is all there is to

1 being phenomenally consciousness according to the HOROR theory.<sup>6</sup>  
2 When one has a HOROR like this, one will be phenomenally con-  
3 scious and will experience the blue sky, the ripe tomato, the bright-  
4 tone trumpet, or whatever it is that such a HOROR says you are  
5 currently representing. Notice that this is not to say that the world *is* as  
6 we experience it! But the important point is that HOROR theory  
7 vindicates the common-sense idea that in consciously experiencing the  
8 blueness of the sky, the redness of the tomato, or the brightness-tone  
9 of the trumpet, it does not necessarily seem to me as though I am  
10 aware of some mental quality. I seem to be aware of the tomato itself  
11 and its redness. According to the HOROR theory, this is because the  
12 higher-order representations describe oneself as being in states which  
13 present one with objects in the environment that have colours, sounds,  
14 etc. and so this is how one experiences the world.

#### 15 **4. Higher-Order Representation Intentionally** 16 **For Introspective Consciousness (HORIFIC)**

17 What is the HORIFIC theory? When one has a conscious experience  
18 of seeing blue and one then introspects such a mental state, one still is  
19 visually experiencing blue but one is also aware of the blueness as a  
20 property of one's visual experience. 'What it is like' for one is *like*  
21 consciously probing and focusing to one's own mental life. When one  
22 introspects a conscious experience, one is in a special relationship  
23 with one's conscious experience.<sup>7</sup> A brief comparison is worthwhile:  
24 While HOROR is a theory of phenomenal consciousness which  
25 accounts for the character of the experience as a property of — or  
26 related to — an object of the external world, the HORIFIC theory  
27 extends to introspective consciousness and builds upon the HOROR  
28 theory. The HORIFIC theory accounts for the distinctive way in

---

<sup>6</sup> One reason for adopting HOROR theory comes from Block's (2011) attack on Rosenthal's HOT theory. If there is phenomenal consciousness in the absence of the *first-order state*, then we seem to have a conscious state with no neural correlate. Rosenthal seems to agree that the correlate of the conscious experience is the correlate of the *higher-order representation*. So, we see him as ultimately agreeing with us about HOROR. Further discussion and analysis must await another occasion.

<sup>7</sup> Advocates of the transparency of experience may deny these claims about introspection; defending them is beyond the scope of this article. Here, we note that the kind of examples presented at the beginning of the paper involving taking off one's glasses suggest that one can become aware of one's experience. This suggests that the strong claim that we are only ever aware of properties of external objects and never aware of any of our mental properties is false.

1 which the relevant experience appears *to oneself* as a property of — or  
 2 related to — one’s own experience or occurrent mental state.

3 We can put the HORIFIC theory of introspective consciousness  
 4 more specifically:

5 **HORIFIC:** For a subject *S* to be introspectively conscious of a  
 6 mental state *m* as having character *Q* is for *S* to token an appro-  
 7 priate higher-order representation with content that *S* is experi-  
 8 encing *m* with character *Q*.

9 The HORIFIC theory, then, accounts for the distinctive way in which  
 10 a first-order representation is conscious to the subject. A caveat here is  
 11 important. For both HOROR and HORIFIC theories, the distinctive  
 12 ‘what-it-is-like-ness’ of an experience is exhausted by the representa-  
 13 tional content of the appropriate higher-order representation. How-  
 14 ever, different versions of the HORIFIC theory may postulate  
 15 different contents depending on how they characterize the term  
 16 introspective consciousness. As for now, we have not developed other  
 17 versions of the theory.

18 Suppose one is looking at a basket of ripe strawberries which are  
 19 deeply red and looking delicious. As we have seen, according to the  
 20 HOROR theory one’s experience of the redness and fragrance, for  
 21 example, consists entirely in tokening an appropriate higher-order  
 22 representation attributing to oneself being in states which represent the  
 23 strawberries. But what happens when one becomes introspectively  
 24 conscious? According to the HORIFIC view, one comes to form an  
 25 introspective HOROR. That is, one comes to form an introspective  
 26 representation of the same first-order representations one was already  
 27 targeting. The targeted first-order representations are themselves  
 28 mental states which represent the strawberries — and which have their  
 29 own functional roles. These states may involve mental qualities,  
 30 concepts, or whatever it is that one thinks allows the creature in  
 31 question to be aware of the properties in the environment that matter  
 32 to it. Let us call this collection of first-order representations *m* and the  
 33 original HOROR *C* (for consciousness). Then, *C* targets *m* and repre-  
 34 sents oneself as being in *m*. When one then introspects, *C* can come to  
 35 be replaced by *I* — i.e. an introspective HOROR — where *I* targets  
 36 the very same first-order representations *m* and represents oneself as  
 37 experiencing being in them.

38 The introspective HOROR is *not* an additional higher-order repre-  
 39 sentation in the sense of a third-order representation of the original  
 40 HOROR *C* from above. Rather, it is just a *different* second-order

1 representation that targets the very same first-order perceptions but  
2 with a new introspective content. Introspective consciousness is a kind  
3 of phenomenal consciousness — a probed and focused version of  
4 ordinary non-introspective consciousness — and the HORIFIC view  
5 explains it in the same way as it does phenomenal consciousness  
6 generally.

7 On the HORIFIC view, the higher-order representation represents  
8 oneself as experiencing the first-order representations. The content of  
9 this introspective higher-order representation, roughly, would be that *I*  
10 *am experiencing — perceptible — red in a distinctly visual way*. Since  
11 this introspective representation is a HOROR, it follows (*ex hypothesi*)  
12 that the subject has an experience which is exhaustively characterized  
13 by the representational content of the HOROR. In this case one would  
14 be *experiencing seeing red*. This is what it is to introspect the redness  
15 of the strawberries on the HORIFIC view.

## 16 5. HORIFIC: A Pluralist View about 17 Modes of Introspective Consciousness

18 If the HORIFIC theory is true, there is room to account *both* for two  
19 different modes of introspective consciousness *and* for the nature of  
20 the first-order representation at issue — or at least certain properties  
21 of the target mental state.

22 As for the former claim, if introspective consciousness is considered  
23 of two modes, this will open up the possibility of postulating a  
24 pluralist approach, either:

25 **Stimuli-induced:** A receptive introspective consciousness  
26 whereby a specific mental state or first-order state spontaneously  
27 or automatically induces a higher-order representation (see  
28 Renero, 2019, p. 837).

29 **Self-triggered:** A selective introspective consciousness whereby  
30 the subject's own interest and volition initiate a higher-order  
31 representation of an occurrent first-order representation (*ibid.*, see  
32 p. 840).<sup>8</sup>

---

<sup>8</sup> It is important to clarify that we do not mean here having any voluntary control over the mental state or a capacity to inhibit the higher-order representation at issue. But we do refer to being able to be introspectively conscious of what is there to be represented. In this sense, there is an active role being undertaken by the subject in this particular mode of introspective consciousness.

1 On the one hand, in considering *stimuli-induced* introspective con-  
2 sciousness, it is important to see that the stimulus is not necessarily  
3 caused or induced by an external source of the physical world but is a  
4 (first-order) mental-state-induced representation. Stimuli-induced intro-  
5 spective consciousness shows that the character of the mental stimulus  
6 induces different representations and exhibits distinct outputs, which  
7 can be either simple representations or complex representations that  
8 emerge, such as when you hit me: I respond with pain, I shout or  
9 express anger, I hit you back, or run away.

10 Following our initial case of (a target mental state such as) a pain  
11 sensation, consider an example of a *simple output*: a first-order state  
12 such as a severe pain sensation in your toe induces an introspective  
13 HOROR of that state and you describe such a state as a stabbing: ‘I’m  
14 experiencing having a stabbing sensation’ — after ruling out a  
15 throbbing sensation, for example. Let us call it a higher-order repre-  
16 sentation induced by a painful sensation.

17 Now, consider an example of a *complex output*: a first-order state  
18 such as a severe pain sensation in your toe induces a higher-order  
19 representation of that state, which brings about a new target state: a  
20 visual state of your swollen toe. And this state may also induce a  
21 simultaneous mental state: a deep concern about paying for a medical  
22 consultation to get an examination of your toe, or a real desire that I  
23 — the tormentor who has stomped on you — disappear instantly.  
24 Accordingly, you describe such a state as a swollen toe: ‘I’m experi-  
25 encing having a swollen toe, now I need to see a doctor, I do not have  
26 health insurance and I’m on bankrupt.’

27 On the other hand, in considering *self-triggered* introspective con-  
28 sciousness, it is important to see a different scenario. First, take the  
29 following context: as always, you are rushing. But you prefer to stop  
30 for a minute to enquire as to what your current first-order state is. You  
31 choose to start a self-probing of your own mental life and form a  
32 higher-order representation of your current mental state — i.e. one  
33 first-order state that you select from a cluster of occurring first-order  
34 states as they appear in your mind. Now, consider the previous pain  
35 sensation example as happening to you right now: you form a higher-  
36 order representation of your current pain sensation — and set aside for  
37 the moment other first-order states that you can generate — e.g. any  
38 emotional response to it, or your desire to take revenge. Self-triggered  
39 introspective consciousness is accompanied by an intention to under-  
40 take a search, or by simple inquisitiveness as one self-probes one’s  
41 first-order mental representation for the sake of investigation.

1 As mentioned, adopting the possibility to account for distinct modes  
2 of introspective consciousness — stimuli-induced or self-triggered —  
3 promotes a pluralist approach. The HORIFIC theory is pluralist since  
4 introspective consciousness would not be restricted to a single  
5 relation, nor would it be reducible to a unique form of access or  
6 awareness of one’s mental states.<sup>9</sup> Instead, we propose that higher-  
7 order representations can represent first-order mental states in distinct  
8 modes. Those modes may vary depending on time and specific  
9 situation.<sup>10</sup> While stimuli-induced introspection may exhibit distinct  
10 types of outputs, self-triggered introspection may reveal the selected  
11 mental state in various ways and furnishes the experience. A pluralist  
12 approach such as this allows the possibility of broadening the notion  
13 of introspective consciousness while other theories of introspection  
14 either neglect it or maintain constraints.<sup>11</sup> We will not engage here in  
15 an examination of these theories, or in a comparison or contrast  
16 between HORIFIC and other approaches.<sup>12</sup> The relevant point here is  
17 that this novel consideration had also been left out by leading views of  
18 consciousness.<sup>13</sup>

---

<sup>9</sup> One can also be a pluralist concerning the variety of the target mental states that are represented via HORORS. That is, we can have higher-order representations that target *all* types of target mental states — i.e. *cognitive states* or propositional attitudes such as beliefs and *non-cognitive states* such as sensations and perceptions or those experiences that are linked to our sensory modalities (seeing, hearing, smelling, tasting, and touching); and our affective states or those experiences such as emotions, feelings, and bodily sensations like pains, itches, tingles, and cramps — or we can have higher-order representations that target only *some* of them. Although we think that HORIFIC can represent *all* kinds of states as its *targets*, offering an account of this possibility goes beyond the scope of this paper.

<sup>10</sup> Notice that an introspective episode can begin by self-triggered introspective consciousness and end up with a verbal report, or it can begin by stimuli-induced introspective consciousness and end up solely in a silent judgment.

<sup>11</sup> For example, single theories such as the *inner-sense* view account for a mere causal or mediated relation between the target mental state and the introspection of such a target. Whereas theories such as the *acquaintance* view account for a direct or immediate relation between the target mental state and introspection of such a target. These theories claim that only one type of awareness, either the causal one or the direct one, is defined as introspection. Presenting a non-exclusive alternative view which can accommodate both relations may be more attractive than other restrictive positions. For a discussion between these views and a possible conciliation, see Renero (2019).

<sup>12</sup> Exploring whether the HORIFIC theory is compatible with the mentioned views requires a different analysis that goes beyond the present purview.

<sup>13</sup> For example, HOT theory is usually considered a *single* theory — as opposed to a pluralist theory — in which the higher-order thought is caused by a stimuli of the external world. HOT theory is also identified among the inner-sense views similar to

1        Although the fundamental differences between stimuli-induced and  
 2 a self-triggered introspective consciousness are clearly seen in the  
 3 definitions and characterizations above, close attention to the previous  
 4 scenarios reveals that introspective consciousness — in either form —  
 5 fulfils three criteria, at least: ‘it is directed at one’s mind (first-person);  
 6 it is about psychological states; namely, mental entities, as opposed to  
 7 non-mental entities (mental); and it is about one’s current, ongoing,  
 8 and recent past mental states (occurring)’ (Renero, 2019, p. 824).  
 9 Now, introspective individuation of one’s occurring states or experi-  
 10 ences is given as a function of these criteria plus other conditions of  
 11 access such as consideration of context, possession of right cognitive  
 12 capacities, being able to offer a justification of an occurring or  
 13 ongoing mental state, and having the relevant conceptual resources to  
 14 describe or report one’s own mental states (*ibid.*, see p. 833). These  
 15 criteria are presupposed in the examples that we have offered here.<sup>14</sup>

16        Let us turn to the latter claim — i.e. that the HORIFIC theory can  
 17 also tell us something about the nature of the first-order target mental  
 18 state at issue. That is, the HORIFIC view can account *both* for which  
 19 specific mental state is the target of a particular higher-order repre-  
 20 sentation — e.g. whether my first-order representation is a pain  
 21 sensation versus a tickle, or whether it is a desire rather than a whim  
 22 given its particular phenomenal character — *and* for the content or the  
 23 way in which one represents oneself as having the experience and how  
 24 one can eventually describe the relevant first-order state in terms of  
 25 *changes*, *transitions*, and *boundaries* between that state and other  
 26 occurring states in accordance with time and specific situation.

27        As for the nature of the first-order representation or certain  
 28 properties of the target mental state, we contend that:

- 29        P1. Subject *S* can introspectively individuate mental states according  
 30 to their properties as changing or *varying*, transitioning or  
 31 *passing through*, and dividing or *delimitating*, while they are  
 32 appearing to *S*’s mind and *S* forms a corresponding higher-order  
 33 representation.

---

Armstrong’s view (1968/1993). Discussion about this particular issue is left for future work.

<sup>14</sup> Although we acknowledge that offering analysis on the metaphysics of experiences is relevant and necessary for further justification, we will have to leave this analysis for future work.

1 P2. If P1 holds, the HORIFIC theory accounts for introspective con-  
 2 sciousness as a dynamic and flexible phenomenon and HORIFIC  
 3 can be a promising theory.

4 \_\_\_\_\_

5 C. The HORIFIC theory accounts for introspective consciousness as  
 6 a dynamic and flexible phenomenon and HORIFIC can be a  
 7 promising theory.

8 First, concerning *changes*, consider (a) the change or shift of direction  
 9 from a recently past mental state to a current mental state within a  
 10 specific introspectively conscious event ( $ms_1$  at  $t^1$ ,  $ms_2$  at  $t^2$ ), or (b) the  
 11 *change* — which can be sudden or gradual — in phenomenal  
 12 character of an experience within a specific introspectively conscious  
 13 event ( $e_p$  at  $t^1$ ,  $e_p$  at  $t^2$ ). For example, (a) while you are forming a  
 14 higher-order representation of your visual perception of a bottle of  
 15 mezcal, from the phenomenal character of this visual experience you  
 16 can be introspectively conscious of a sudden change in your target  
 17 mental state: your visual perception has just shifted into the desire to  
 18 have a glass — or two! — of mezcal. Notice here that the type of  
 19 target mental state at  $t^1$  changes to another target mental state at  $t^2$ .  
 20 Alternatively, (b) while you are probing your gustatory sensation of  
 21 the mezcal and enjoying its strong smoky flavour, you find out by  
 22 introspective consciousness that the character of your experience  
 23 suddenly changes when you see a worm inside the bottle. Notice here  
 24 that the phenomenal character of the target mental state occurring at  $t^1$   
 25 changes given a new content of the representation at  $t^2$ .

26 Second, concerning *transitions*, consider the transition from an  
 27 occurring mental state to a fading or dissipating mental state as a  
 28 specific introspectively conscious event (from  $ms_1$  to  $ms_0$  at  $t^1$ ), or the  
 29 pathway from having a powerful experience to including it within an  
 30 introspectively conscious event (from  $e_1$  to  $e_0$  at  $t^1$ ). For example,  
 31 while you are probing your auditory state of hearing yourself laugh  
 32 and enjoying the bodily sensation that it brings about, you detect by  
 33 introspective consciousness the transition in the state from a chuckle  
 34 to a smile, to a feeling of embarrassment; and all this precisely at the  
 35 time you notice you are laughing with a horrible screeching sound.

36 Third, concerning *boundaries* which divide or delimitate one mental  
 37 state from another according to its duration or scope,<sup>15</sup> consider the

---

<sup>15</sup> Even if the introspective event is relatively short, its duration is susceptible to further extension if introspective consciousness continues as it might (or might not).

1 boundary or the limit between (a) two joint mental states occurring at  
 2 the same time within a specific introspectively conscious event ( $ms_1$   
 3 and  $ms_1^*$  at  $t^1$ ), or (b) two similar or phenomenally related mental  
 4 states having different content within a specific introspective event  
 5 ( $ms_1$  about  $s$  and  $ms_1$  about  $d$ ). For example, (a) while you are probing  
 6 your auditory sensation, which you savour, of simultaneously (audi-  
 7 ble) crying and laughing, you detect by introspective consciousness  
 8 the boundaries between this particular auditory experience — usually  
 9 characterized by a series of spontaneous sounds — and the delightful  
 10 experience that accompanies it. Although these states occur at the  
 11 same time, they have different duration or scope. Either that, or  
 12 (b) while you are self-probing your mental state of sadness because of  
 13 a particular loss, you discover by introspective consciousness the  
 14 boundary between this sadness expressed by crying is, on the one hand  
 15 — or in the beginning — represented as sorrowful, and, on the other  
 16 hand — or when it is coming to a close — represented as desperate.

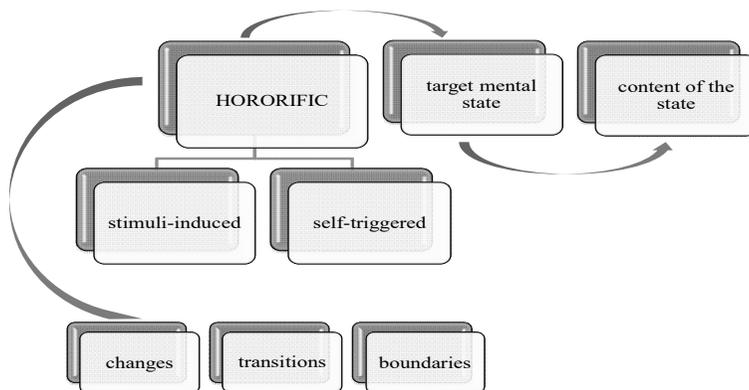
17 Boundaries are particularly helpful in recognizing how instances of  
 18 target mental states of shortened duration occur, how repetitions of a  
 19 certain mental state come along, and also in delimitating when exactly  
 20 the target mental state starts and ends, when an earlier mental state is  
 21 similar to a current state, and so forth. Thus, we say that one is able to  
 22 form corresponding higher-order representations of one as having  
 23 those particular experiences. The formation or generation of higher-  
 24 order representations according to changes, transitions, and bounda-  
 25 ries is given as a function of their corresponding first-order states —  
 26 i.e. in how many distinctions in her conscious experiences the intro-  
 27 spective subject can attain. So, we say that the introspective subject  
 28 will be having higher-order representations that assert that she is in a  
 29 particular mental state with certain phenomenal character and content.  
 30 Differences in phenomenal character are reflected in the intentional  
 31 content of her representations, which are themselves more refined in  
 32 descriptions or verbal reports.

33 It is relevant for our theory to emphasize that the introspective  
 34 subject is aware of being in certain first-order mental states as they  
 35 change sequentially, the transition between themselves, or the bounda-  
 36 ries they maintain. These introspective experiences, on HORIFIC,  
 37 amount to the changing contents of the relevant higher-order states,  
 38 which themselves are driven or partially driven by the causal  
 39 connections to the first-order states themselves (stimuli-induced) and/  
 40 or initiated or partially initiated by the own interest and volition of the  
 41 subject who is introspecting (self-triggered). Those higher-order states

1 can be combined such as when introspective consciousness is driven  
 2 by a stimulus and, then, the output demands self-triggered or further  
 3 introspective consciousness about the target mental state.<sup>16</sup> These con-  
 4 siderations involve the possibility of accounting for higher-order  
 5 representations in accordance with the mentioned two different modes  
 6 and depending on the property or the content of the first-order state to  
 7 be represented.

8 On the HORIFIC view, introspective consciousness is a flexible and  
 9 dynamic phenomenon; it can respond to the changes of one's experi-  
 10 ences according to their phenomenal character and content, while identi-  
 11 fying and distinguishing other accompanying mental states. Introspect-  
 12 ive consciousness is also responsive to fluctuating mental stimuli by  
 13 its flexibility to switch from one mental state at  $t^1$  to another mental  
 14 state at  $t^2$  and, accordingly, to form different higher-order representa-  
 15 tions. The same applies to transitions and boundaries between the  
 16 mental states that accompany other higher-order representations.

17 To close this section, let us illustrate the core of the HORIFIC view  
 18 to show what this theory accounts for in the following diagram  
 19 (details and examples have been offered above).



20  
 21 *Figure 1.* Introspective consciousness can be receptive (*stimuli-induced*) or  
 22 selective (*self-triggered*). A higher-order representation for introspective  
 23 consciousness can reveal the mental state which is represented (the *target*  
 24 state) and the way in which the target state is represented (the *content* of  
 25 such a state). It can also distinguish *changes*, *transitions*, and *boundaries*  
 26 between mental states in certain episodes of introspective consciousness.

<sup>16</sup> Whether both modes of introspective consciousness can occur together regarding the same mental state is a subject that remains to be worked out in future research.

1

## 6. Conclusion

2 We have offered the basics of the HOROR theory of phenomenal con-  
3 sciousness and have built upon it to extend it to introspective con-  
4 sciousness via the grounds of — what we have called — the HORIFIC  
5 theory. The relevant point is that one would be able to form a higher-  
6 order representation of a first-order representation — i.e. the target  
7 mental state — and thus be introspectively conscious of such a repre-  
8 sentation. Having highlighted the merits of HORIFIC by demonst-  
9 rating this view at work, we have further shown that some of the  
10 theoretical consequences of holding this novel view entail the possi-  
11 bility of providing a pluralist perspective: the target mental states of  
12 higher-order representations are not formed just in one way or repre-  
13 sented in the same way. Rather, one can represent first-order states  
14 and certain properties of those states based on one’s own experience  
15 of certain stimuli, or one can represent first-order states and certain  
16 properties of those states based on one’s own interest and volition to  
17 initiate a self-probing process. This is a promising theory for investi-  
18 gating the relevance of higher-order approaches in connection to intro-  
19 spection, while paving the way for further research on introspective  
20 consciousness.

21

## References

- 22 Armstrong, D.M. (1968/1993) *A Materialist Theory of the Mind*, New York:  
23 Humanities Press.
- 24 Berger, J. (2017) How things seem to higher-order thought theorists, *Dialogue*, **56**  
25 (3), pp. 503–526.
- 26 Berger, J. & Brown, R. (2021) Conceptualizing consciousness, *Philosophical*  
27 *Psychology*, **34** (5), pp. 637–659.
- 28 Block, N. (1995) On a confusion about a function of consciousness, *Brain and*  
29 *Behavioral Sciences*, **18** (2), pp. 227–247.
- 30 Block, N. (2007) Consciousness, accessibility, and the mesh between psychology  
31 and neuroscience, *Behavioral and Brain Sciences*, **30** (5–6), pp. 481–99;  
32 discussion, pp. 499–548. doi: 10.1017/s0140525x07002786
- 33 Brown, R. (2014) Consciousness doesn’t overflow cognition, *Frontiers in Psychol-*  
34 *ogy*, **5** (1399). doi: 10.3389/fpsyg.2014.01399
- 35 Brown, R. (2015) The HOROR theory of phenomenal consciousness, *Philo-*  
36 *sophical Studies*, **172** (7), pp. 1783–1794.
- 37 Brown, R., Lau, H. & LeDoux, J.E. (2019) Understanding the higher-order  
38 approach to consciousness, *Trends in Cognitive Sciences*, **23** (9), pp. 754–768.
- 39 Brown, R., LeDoux, J.E. & Rosenthal, D.M. (2021) The extra ingredient, *Biology*  
40 *and Philosophy*, **36** (2), pp. 1–4.
- 41 Carruthers, P. & Gennaro, R. (2020) Higher-order theories of consciousness, in  
42 Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*, Fall 2020 Edition,

- 1 [Online], [https://plato.stanford.edu/archives/fall2020/entries/consciousness-](https://plato.stanford.edu/archives/fall2020/entries/consciousness-higher/)  
2 [higher/](https://plato.stanford.edu/archives/fall2020/entries/consciousness-higher/) [22 April 2022].
- 3 Chalmers, D.J. (1999) First-person methods in the science of consciousness,  
4 *Consciousness Bulletin*, DETAILS.
- 5 Chalmers, D.J. (2018) The meta-problem of consciousness, *Journal of Conscious-*  
6 *ness Studies*, **25** (9–10), pp. 6–61.
- 7 Dretske, F. (1995) *Naturalizing the Mind*, Cambridge, MA: Bradford Books/MIT  
8 Press.
- 9 Gennaro, R.J. (1993) Brute experience and the higher-order thought theory of con-  
10 sciousness, *Philosophical Papers*, **22**, pp. 51–69.
- 11 Gennaro, R.J. (1996) *Consciousness and Self-Consciousness: A Defense of the*  
12 *Higher-Order Thought Theory of Consciousness*, Amsterdam: John Benjamins.
- 13 Gennaro, R.J. (2004) Higher-order thoughts, animal consciousness, and misrepresen-  
14 tation: A reply to Carruthers and Levine, in Gennaro, R.J. (ed.) *Higher-*  
15 *Order Theories of Consciousness*, Amsterdam: John Benjamins.
- 16 Gennaro, R.J. (2012) *The Consciousness Paradox*, Cambridge, MA: MIT Press.
- 17 Gertler, B. (2011) *Self-Knowledge*, New York: Routledge.
- 18 Gertler, B. (2012) Renewed acquaintance, in Smithies, D. & Stoljar, D. (eds.)  
19 *Introspection and Consciousness*, New York: Oxford University Press.
- 20 Gottlieb, J. (2019) The collapse argument, *Philosophical Studies*, **176** (1), pp. 1–  
21 20.
- 22 Lau, H. (2019) Consciousness, metacognition, and perceptual reality monitoring,  
23 *PsyArXiv*. doi: 10.31234/osf.io/ckbyf
- 24 Lau, H. & Rosenthal, D. (2011) Empirical support for higher-order theories of  
25 conscious awareness, *Trends in Cognitive Sciences*, **15**, pp. 365–373.
- 26 Lau, H. & Brown. R. (2019) The emperor’s new phenomenology? The empirical  
27 case for conscious experience without first-order representations, in Pautz, A. &  
28 Stoljar, D. (eds.) *Blockheads! Essays on Ned Block’s Philosophy of Mind and*  
29 *Consciousness*, Cambridge, MA: MIT Press.
- 30 Nagel, T. (1974) What is it like to be a bat?, *Philosophical Review*, **83** (4), pp.  
31 435–456.
- 32 Raccah, O., Block, N. & Fox, K. (2021) Does the prefrontal cortex play an  
33 essential role in consciousness? Insights from intracranial electrical stimulation  
34 of the human brain, *The Journal of Neuroscience*, **41** (10), pp. 2076–2087. doi:  
35 10.1523/jneurosci.1141-20.2020
- 36 Renero, A. (2014) Consciousness and mental qualities for auditory sensations,  
37 *Journal of Consciousness Studies*, **21** (9–10), pp. 179–204.
- 38 Renero, A. (2017) The nature of introspection, *CUNY Academic Works*, [Online],  
39 [https://academicworks.cuny.edu/gc\\_etds/2144](https://academicworks.cuny.edu/gc_etds/2144).
- 40 Renero, A. (2019) Modes of introspective access: A pluralist approach, *Philo-*  
41 *sophia*, **47**, pp. 823–844. doi: 10.1007/s11406-018-9989-2
- 42 Rosenthal, D.M. (1986) Two concepts of consciousness, *Philosophical Studies*, **49**,  
43 pp. 329–359.
- 44 Rosenthal, D.M. (1993) Thinking that one thinks, in Davies, M. & Humphreys, G.  
45 (eds.) *Consciousness*, Cambridge, MA: Blackwell.
- 46 Rosenthal, D.M. (1997) A theory of consciousness, in Block, N., Flanagan, O. &  
47 Güzeldere, G. (eds.) *The Nature of Consciousness: Philosophical Debates*,  
48 Cambridge, MA: MIT Press.
- 49 Rosenthal, D.M. (2005) *Consciousness and Mind*, New York: Oxford University  
50 Press.

- 1 Shargel, D. (2016) The insignificance of empty higher-order thoughts, *Journal of*
- 2 *Cognition and Neuroethics*, **4** (1), pp. 113–127.
- 3 Tye, M. (1994) Qualia, content, and the inverted spectrum, *Noûs*, **28**, pp. 159–183.
- 4 doi: 10.2307/2216047
- 5 Paper received November 2021; revised June 2022.