Enacting Virtual Reality: The Philosophy and Cognitive Science of Optimal Virtual Experience

Garri Hovhannisyan[1,4], Anna Henson[2,4(✉)], and Suraj Sood[3,4]

[1] Duquesne University, Pittsburgh, PA, USA
[2] Carnegie Mellon University, Pittsburgh, PA, USA
henson.anna@gmail.com
[3] University of West Georgia, Carrollton, GA, USA
[4] Sirius Project, Melbourne, USA

Abstract. The standard approach to immersive virtual reality (VR) is arguably "object-centric" in that it aims to design physically realistic virtual experiences. This article deems the object-centric approach both philosophically and theoretically problematic and builds up to an alternative, "action-predicated" approach, whose aim is to simulate virtual experiences with a primary emphasis on pragmatic functionality instead. Section 1 lays out the rationale of the article and provides an outline for its general structure. Section 2 illustrates the nature of the problem being tackled and articulates a philosophically motivated critique, demonstrating the necessary limitations of the standard approach, as well as the need for an alternative. Section 3 draws on the enactive approach to cognitive science and begins the formulation of such an alternative. Section 4 completes the turn toward an action-predicated approach and argues, in particular, for a flow-based conception of immersive VR experience. Section 5 systematically discusses the methodological implications of the theoretical merits of this article by examining a design probe, *Wake*, conducted on participants ($N = 25$) in a mixed reality (MR) setting. Finally, Sect. 6 constitutes the conclusion of this article, wherein its philosophical, theoretical, and methodological efforts, as well as possible avenues for future research, are briefly noted.

Keywords: Enactivism · Virtual reality (VR) · Flow · Immersion · Social presence

1 Introduction

This article takes as its point of departure the belief that the contemporary literature on virtual reality (VR) is theoretically confused. In particular, our claim is that VR has inherited and remains tacitly committed to a philosophically impoverished, and hence theoretically implausible, view of human perceptual reality, as a consequence of which its attempts at simulating what is "real" either fall short in practice, or else remain ill-conceived theoretically even despite their apparent practical success. We believe, though, that this confusion, along with its practical challenges, can be remedied, and that the means to remedying it require explicit philosophical engagement with the existing body of literature on VR. We are quite aware that philosophical concerns are

often regarded as irrelevant or only peripheral in importance to matters of technological design and innovation in such fields as VR. However, what we hope to demonstrate with our contribution is an alternative to this view. Specifically, we aim to show that determining how exactly to render virtual experiences more subjectively real or plausible is not only a matter of technological design, but, indeed, also of philosophical and theoretical commitment.

Having presented the overarching rationale of our project, the structure of this article is as follows. In Sect. 2, we articulate a critique of the received view in VR, which sets as the design goal for VR the maximization of "presence". Presence is a perceptual illusion that is brought about when the contents of an immersive virtual simulation are found to be sufficiently "believable" so as to be treated as "real" by the user. We show that this design goal is predicated, first, on the assumption that the reality to be simulated by VR is physical reality (reality as a mere "collection of objects"), and second, that perceptual reality as such is realized strictly in the brain, as apart from or independent of various subjective factors, such as emotion, motivation, and cognition, and without the need for embodied interaction with the world. In our critique, we demonstrate how both of these assumptions are actually rooted in an outdated and particularly reductive brand of physicalist philosophy, such that we find it difficult to imagine how any set of design guidelines for VR that is predicated on these assumptions can help to explain or even generate meaningful virtual experience. We close our critique with the claim that adopting a physicalist approach to VR in this way results in three theoretical "gaps" or challenges, which, we maintain, cannot be bridged by physicalist means alone.

Sections 3 and 4, accordingly, constitute the constructive elements of our contribution. We thus turn to the "enactive approach" in cognitive science, which we believe offers a viable alternative to the physicalism implicit in the received view. For despite undertaking a study of subjectivity that is scientific in kind, enactivism nevertheless manages to maintain a commitment to non-reductionism in so doing. Thus, the enactive approach promises to bridge the three theoretical gaps besetting VR, insofar as it goes beyond what the reductive brand of physicalism is able to offer.

In Sect. 3, we synthesize a theory of perception on enactive grounds, wherein perception is conceived of as a necessarily embodied, interactive, affective, motivated, and cognitive process. In advancing a non-reductive account of perception that is action-predicated (pragmatic), rather than object-centric (physicalist), we demonstrate how the first two theoretical gaps can be bridged. In Sect. 4, we then extend our theoretical framework to also bridge the third theoretical gap. We begin by reviewing recent cognitive scientific work on "the flow state", often described as the state of optimal experience [5, 17], and we integrate these findings with our own enactive theory of perception. We thus explain the phenomenon of subjective immersion in terms of flow, that is, as an action-predicated process of optimal perceptual engagement with the world. We then use our flow-based approach to subjective immersion to explain the phenomenon (and phenomenology) of immersion as it pertains to VR experience (optimal virtual experience). Finally, we conclude this section by arguing that the primary design goal of VR should be, not the maximization of presence through the simulation of physically realistic experiences as mediated by the objectively immersive VR hardware,

but the maximization of subjective immersion by developing virtual experiences that are able to reliably facilitate a flow state within users.

In Sect. 5 we proceed to demonstrate the relevance of our theoretical contribution for praxis. Here, we articulate a set of methodological implications that follow from adopting subjective immersion, rather than presence, as the primary goal for VR design. These implications are grouped in such a way as to address the following four elements of VR experience: (1) "Onboarding", (2) Immersion, (3) "Offboarding", and (4) "Experiential optimization". We conduct our discussion of methodology by examining a design probe, *Wake*, that was conducted with users in a mixed reality (MR) setting, which we use to detail the implementation of the proposed implications in an empirical, real-world setting [18]. Section 6 constitutes the conclusion of this article. In it, we briefly summarize and assess the philosophical, theoretical, and methodological import of our argument and note possible avenues for future research.

2   The Received View and Its Discontents

This section advances the claim that any set of design guidelines that is derived from or predicated upon a strictly physicalist conception of reality can only ever lead to sub-optimal VR experience, when implemented. To this end, we begin by first unpacking what we mean by physicalism and we demonstrate in what way exactly physicalism is assumed within the VR literature. We then proceed to demonstrate the issues associated with assuming physicalism in this way, whereby we identify a total of three theoretical gaps besetting the VR literature, which, we argue, cannot (in principle) be bridged by physicalist means alone. We conclude with the proposition that addressing these gaps, and thus conceptualizing optimal VR experience, requires going beyond mere physicalism. In Sect. 3, we begin our elaboration of this proposition by offering an alternative (non-reductive) conception of perceptual reality that is rooted in the enactive approach to cognitive science.

Physicalism is the view that what is real is, most basically put, just what is physical. In other words, reality consists exclusively of physical entities, such that every observable phenomenon is, in essence, physical in nature, and can ultimately be explained by essentially physical causes. Physicalism is the prevalent position in the natural sciences, and is often assumed as a matter of fact, rather than philosophical commitment. Despite both its prevalence and its success as a doctrine, physicalism is not without its own set of problems—especially when taken for granted within the context of cognitive science and related fields of study (e.g., philosophy of mind, psychology). VR happens to be one such field of study. Even so, the philosophical implications of physicalism for VR have, to this point, remained largely unnoticed, and therefore neglected by theorists. We therefore find it crucial to bring these problems to the fore and to explicitly address them, insofar as much of the theoretical labor and design efforts in VR necessarily depend on the philosophical commitments that are made at the outset.

It is important to begin by first pointing out that the physicalism assumed in the mind sciences—and VR, by extension—is not of the same kind as what is assumed within the natural sciences today. Historically, the mind sciences (particularly,

psychology) tried to fashion themselves after the natural sciences in order to legitimize their standing as a scientific discipline in their own right, and they did so by appropriating the basic philosophical and methodological assumptions of the natural sciences [32]. At the time when this appropriation occured, though—(during the late 19th and early 20th centuries)—the natural sciences were still embedded in a Newtonian worldview, according to which the world is essentially a place of material objects (inert masses) whose motion is governed by a set of mechanical laws (e.g., the law of gravity).

Essential to the Newtonian conception of reality were the following two suppositions: (1) "Materialism", according to which reality is essentially material or physical, and (2) "Mechanism", according to which causality is a wholly linear process whereby every physical event can be explained in terms of some temporally antecedent cause [27, 28]. It goes without saying that the natural sciences have advanced beyond Newtonianism in their construal of reality, particularly in light of the various epistemological innovations witnessed over the course of the 20th century, such as quantum mechanics, Einstein's theory of general (and special) relativity, as well as complexity science and information theory The same, however, cannot be said of the mind sciences, which, on the whole, remain fixated on an outdated (Newtonian) worldview.

In having committed to such suppositions as materialism and mechanism, the general approach to studying the mind, thereof, has been largely reductive. In particular, the materialist supposition has motivated a conception of mental functioning as wholly physical and objective in kind, thus leaving unexplained various core (subjective) properties of the mind, such as consciousness [19], normativity [14, 15], and purposiveness (goal-directedness) [33]. Accordingly, the mechanistic supposition has motivated a conception of mental functioning as a form of information-processing, as exemplified by the metaphor of the mind as a computer most notably used in the discipline of cognitive psychology. According to the information-processing model, the mind is to the software of a computer as the brain is to the hardware. The brain receives information through the senses (input) and processes it in such a way as to produce functional action in the world (output). Cognition, in other words, begins at the moment when sensory information is received, and ends when a functional motor output is generated. Importantly, though, the essential hardware of the mind is just the brain, which is the center of information processing (it is causally primary, and sufficient, for cognition). The information-processing model of the mind is exemplary of the physicalist approach and has stood as the prevalent metaphor in the mind sciences since the 1950's [34].

The summary of the physicalist grounding motivating the mind sciences provided here is not meant to constitute an exhaustive account, but only a rough sketch. Our current goal, after all, is to articulate an understanding of how some of the fundamental presuppositions within the mind sciences have (mis)informed, and continue to (mis)inform, VR research and design. Our claim here is that VR has tacitly inherited a physicalist (Newtonian) conception of reality as conceived as a collection of material objects, as a consequence of which both VR research and design have operated under a philosophically impoverished theory of mind and subjectivity. We find particularly problematic and worthy of discussion three implications of physicalism for VR—in other words, three theoretical gaps—to which we now turn.

2.1 The Three Theoretical Gaps

To begin with, the fundamental physicalist claim assumes that perception of reality is essentially just the perception of physical objects in an a surrounding space. Once appropriated by VR, what follows from this claim is that perception within virtual reality should simply occur as a matter of course, if only the sensory information that is simulated within the virtual environment is sufficiently high in fidelity to the sensory information that is naturally present in ordinary, physical reality. Technically speaking, therefore, VR subscribes to a "naïve realist" theory of perception, according to which perceiving reality is ultimately a passive process, insofar as reality is construed as essentially physical, objective (mind-independent), and therefore ontologically "pre-given" and "readymade" for perception. Naïve realism has been heavily criticized by a wide range of disciplines, such as philosophy (Descartes, Hume, Kant), science and technology studies [26], and social theory [1]. Repeating any of the arguments made by these authors would at this point be too redundant a step to take. Thus, we will confine ourselves to but one example to help illustrate our point.

VR experience typically entails a period of adjustment or recalibration in the beginning, which we will refer to as an "onboarding" process. During onboarding, the participant's perceptual systems are preoccupied with realizing an adaptive fit with the VR interface and its concomitant hardware. Indeed, what is entailed by onboarding is not just an appropriation of the contents of perception in VR, but also—and perhaps even primarily—an appropriation of the medium of perception of VR (i.e., the VR hardware and its attendant interface for control). Particularly illustrative of this point is the fact that, upon being equipped, the VR headset initially acts like a blindfold, insofar as it severs the individual's visual connection with the world. Upon being "blinded" in this way, what often becomes most salient in the participant's awareness is the various physical qualities (e.g., weight, temperature, texture) of the headset as it is felt against one's face and head. However, as the individual spends time interacting with the immersive virtual environment, the "headset-as-blindfold" eventually becomes a "headset-as-window-into-the-virtual-world", such that the physical qualities of the headset, as well as those of the rest of the augmented interface, recede into the background of the individual's awareness. As the physical properties of the VR interface become less salient, a more stable perceptual connection is formed between the individual and the VR content. In other words, through the onboarding process, the individual's attention gradually shifts from being focused on the medium of perception (i.e., the VR hardware and interface) to the contents of perception (i.e., the VR environment) [16], as a result of which the individual's felt perceptual engagement with the VR becomes increasingly "natural" or "intuitive".

The key point here is that perception in VR (normally) does not begin by feeling intuitive, but becomes increasingly more intuitive over time and with continuous engagement with the VR content. This pattern of development, however, is precisely the opposite of what a naïve realist theory of perception would predict, according to which perception would occur passively and instantaneously, rather than progressively. The first theoretical gap within the literature may therefore be phrased as follows: Perception is not fundamentally a passive process, contrary to the naïve realism implicit in VR, but is necessarily dependent upon one's voluntary patterns of participation in

and interaction with an environment—be it virtual or ordinary. This is to say that perception is more accurately construed as a process of "skillful coping"—to borrow a term from Dreyfus [20]—one that must be learned and achieved as a matter of continuous sensorimotor coordination in relation to, and mastery over, a meaningful situation; rather than as a matter of merely "receiving" sensory information from the environment and processing it instantly and strictly inside one's own "head". A naïve realist theory of perception that is motivated on physicalist grounds simply cannot account for the fact that perception in VR becomes second-nature only with due diligence, rather than instantly and matter-of-factly, right at the outset of the VR experience. Nor can it account for the fact that an individual's perceptual engagement with the VR content appears to be predicated on skill, in that it is learned by way of active, voluntary, and embodied interaction with the VR content. This interaction is a constant process of negotiation with evolving environmental, phenomenological, and social factors at hand. In other words, this is an embodied interaction, in line with the theory articulated by Dourish [21].

The second theoretical gap implicated by physicalism, accordingly, is predicated on the assumption that factors such as affect, value, thought, as well as motivation, in virtue of their being characteristically subjective (mental) properties, are construed neither as part of reality itself, nor as amenable to scientific theorization. Both of these implications, however, are dubious. For starters, the claim that subjective factors are not part of reality is evidently physicalist in kind, meaning that it is not philosophically neutral (insofar as physicalism is not a philosophically neutral position) and should therefore not be simply taken for granted. Furthermore, the preclusion of subjective factors from reality generates a rather peculiar tension within VR. On the one hand, one of the overarching aims of VR is to develop meaningful experiences by way of immersion. It goes without saying, though, that meaningful experiences, in virtue of their being a type of *experience*, are (at least) partly subjective in their constitution and should therefore be accounted for, rather than precluded, by a theory of perception. However, because of its physicalist bias, VR sets out to simulate reality just as a place of objects, propounding as a result a rather partial and ontologically impoverished view of perceptual reality (naïve realism). In our view, which we elaborate more fully in Sect. 3, the objects of perceptual reality are irreducible to the objects of physical reality, such that any attempt to generate meaningful virtual experience can proceed successfully only by adopting a non-reductive stance with regards to the nature of perceptual reality. We believe, in other words, that a sustained commitment to a naïve realist theory of perception that is predicated on physicalism will only hinder VR's goal of generating meaningful (and optimal) virtual experiences, and should therefore be replaced with a theory of perception that is neither naïve realist nor Newtonian in its ontological commitments.

As regards with the second implication, despite the historical difficulties and resistance associated with subjecting factors such as affect, value, thought, and motivation, to scientific (empirical) investigation and scrutiny, it is far from the case that such characteristically subjective factors are not amenable to scientific theorization by today's standards. Quite the opposite: There has been an explosion of scientific research in the literature on just the subjective dimension of human experience [12, 13, 34, 35, 36]. As a consequence, there is not only good reason to believe that it is

possible to theorize about how subjective factors, such as the aforementioned, interact with objective (environmental) factors in a scientifically rigorous manner, but also that perception of physical reality is powerfully and unconsciously motivated by such subjective factors and cannot be understood apart from them, but only as an abstraction devoid of any real-world meaning. In spite of the various advances made in studying subjectivity, particularly within the cognitive sciences, VR has yet to seriously engage with the relevant literature by which to bridge its second theoretical gap, which concerns the nature of the relationship of objective and subjective factors that are involved in perceptual experience.

The final theoretical gap worth noting concerns some pervasive confusions associated with two core constructs within the VR literature: Presence and immersion. Presence is defined as "the subjective experience of being in one place or environment, even when one is physically situated in another" [7, 10] and is typically treated as the golden standard or main measure for how real or "successful" a virtual experience can be said to be. In other words, presence is the current design goal for VR [3], meaning that improving virtual experience (making it feel more real) is primarily a matter of maximizing presence. Various authors (e.g., [3, 4, 9]) have further articulated the presence construct to apply also to such aspects of VR as one's experience of one's own virtual body (self-presence) as well as that of other agents (social/co-presence). Definitions of immersion, on the other hand, have been more variable. Specifically, immersion generally has been defined as either objective or subjective in kind [8]. Objectively, immersion has been regarded as a function of the VR hardware, depending, for instance, on the number of built-in sensors. Subjective definitions, on the other hand, refer to immersion as a state or feeling of "being caught up in another world" [8]. Of course, the problem with the objective definition is that it is not necessarily predictive of subjective feelings of immersion, insofar as it is possible either to feel subjectively immersed in an objectively non-immersive game, like Tetris, or else to lack a subjective sense of immersion altogether when inside of an objectively immersive simulation (as in the onboarding process). Conversely, the problem with the subjective definition of immersion is that it is too similar to definitions of presence so as to lend itself rather easily to conflation fallacies ([8], p. 1409).

We believe it to be possible to circumvent the conceptual difficulties associated with the immersion construct by adopting a flow-based definition of immersion [8]. Flow is a widely researched phenomenon within psychological science with both objective (performance) and subjective (phenomenological) measures, and is commonly referred to as the state of optimal experience [5]. Flow is considered an optimal state for several reasons. First, it occurs when individuals are engaged in challenging tasks, but only when the difficulty level of the task is just beyond the individual's own level of competence in dealing with the task [17]. Phenomenological descriptions of flow often involve a loss of a sense of time, a reduction in levels of reflective self-consciousness, and high levels of immersion in the task at hand. Moreover, feelings of immersion in flow are typically accompanied with the ironic sense of the immersion being effortless (despite the high demands of the task), whereby one feels as though one is truly "flowing" through the experience. Importantly, flow is experienced independent of age, sex, gender, culture, or language. Flow is, in other words, a human universal [5] and is therefore, in a deep sense, a constituent element of human experience. It is also

predictive of psychological well-being, as well as general life satisfaction, and therefore stands as a psychological theory of optimally meaningful experience, as such [29].

In defining subjective immersion in VR as an "experience of feeling totally involved in and absorbed by the activities conducted in a [virtual] place or environment", it thus becomes possible to clearly distinguish the pragmatic, task-related elements of VR experience from the objective elements in the virtual environment (e.g., the surrounding space, social agents, and self), with which presence is chiefly concerned. The implications of adopting a flow-based definition of immersion in this way are twofold. First, in light of the fact that flow describes optimal experience within ordinary reality, it follows quite straightforwardly that immersion-as-flow should therefore describe optimal experience within *virtual* reality. In other words, by adopting a flow-based definition of immersion, the design imperative of maximizing presence by developing physically realistic virtual experiences becomes only secondary in importance to that of maximizing *immersion* by developing virtual experiences that are able to reliably facilitate a flow state within participants. The criterion of realness as it pertains to VR, therefore, fundamentally becomes predicated on immersion, rather than presence. Second, it should become evident that a naïve realist theory of perception can no longer be used as a guidepost for developing VR experiences, insofar as it cannot account for immersion-as-flow in perceptual terms. This is because naïve realism assumes, as per its physicalist (Newtonian) heritage, that reality is ultimately a place of material objects; whereas flow—and by extension immersion—is fundamentally a process. As such, the final theoretical gap can be elegantly summarized as follows: Simulating an environment conducive to flow is a matter of simulating a process, whereas naïve realism can only describe or be used to simulate objects; therefore, the naïve realist theory of perception implicit in VR must be replaced with an alternative theory of perception such that can account for flow—and, therefore, immersion—in processual terms.

In light of the critiques we have leveled in this section, our preliminary conclusion is that the naïve realist theory of perception assumed in VR, as well as the (Newtonian) physicalist grounding upon which it is motivated, cannot be utilized for developing or explaining optimal virtual experience. We have tried to justify this conclusion by illustrating three theoretical gaps which, we have argued, result as a necessary consequence of VR's tacit philosophical commitment to physicalism. The first gap concerns the difficulty in accounting for the skillful nature of perceptual engagement in naïve realist terms, as evinced by the gradual and progressive, rather than instantaneous, development of perception involved in the onboarding process. The second gap concerns the difficulty in conceptualizing perceptual reality scientifically without also reducing the objects of perceptual reality to those of physical reality. Finally, the third gap concerns the difficulty in accounting for immersion, taken as an instantiation of the flow state within VR, in processual rather than objective terms.

The problems posed by these three theoretical gaps are irresolvable by physicalist (Newtonian) means alone, since physicalism lays at their very foundation. We believe, however, that the situation can be remedied, and, furthermore, that the theoretical and philosophical, as well as psychological, resources for remedying it are readily available within the burgeoning field of "enactivism" [12–15, 35, 36]. Enactivism advocates a non-reductive, pragmatically grounded, and fundamentally embodied approach to conceptualizing mind and experience. In the following section, we draw on the

enactive approach to cognitive science to articulate a theory of perception as motivated action, which promises to bridge the first two theoretical gaps, thereby constituting a viable alternative to the naïve realist theory of perception that is assumed in VR. We then synthesize our theoretical findings in Sect. 4 to offer an explanation of flow (and subjective immersion) in processual terms, thereby demonstrating how the third and final theoretical gap can also be bridged.

3   Perception and Enaction

The enactive approach to cognitive science was launched in the early 1990's as an alternative to conceptualizing mind and consciousness to the information-processing model discussed in Sect. 2 [12, 13]. Since then, enactivism has burgeoned into a comprehensive theoretical framework and research programme with an extensive philosophical grounding in traditions such as phenomenological philosophy, philosophical pragmatism, complex and dynamical systems theory, and systems biology [35]. Enactivism promises to offer an account of mental life that is essentially non-reductive (and non-Newtonian), whereby conscious experience is regarded as irreducible to, albeit fundamentally tied with, physical (biological) matter, and wherein cognition is construed in dynamical (non-linear and ecological), rather than mechanistic (linear and self-enclosed) terms.

For enactivism, cognition is ultimately predicated on action and is necessarily bound up by the practical aims of the agent. Moreover, knowledge of how to act (procedural know-how) is taken to be both primary and prior to the kind of knowledge that is concerned with facts and inferences (propositional know-that). In other words, cognition most fundamentally aims to render the world sufficiently predictable so as to afford functional action, rather than being a disembodied (multiply realizable) function of the brain that is primarily aimed at representing reality "as it is". Whereas the physicalist brand of naïve realism inherent in VR depicts reality as a place of ontologically pre-packaged objects, thereby overlooking this pragmatic dimension of perceptual experience, enactivism honors it, and, in so doing, is able to articulate an account of perception that is constitutively motivated, affective, and procedural. The enactive treatment of cognition (and knowledge) as pragmatic is central to the argument we aim to advance in this section, for we believe it constitutes a viable alternative for conceptualizing perception to the physicalist approach assumed in VR. It is thus the aim of this section to demonstrate how an account of perception that is grounded in the enactive approach can be used to bridge the first two theoretical gaps identified in Sect. 2. We then synthesize our findings into a processual account of flow (and immersion) in Sect. 4, thereby demonstrating a bridging of the final theoretical gap.

3.1   Perception as Motivated Action

The world is incomprehensibly complex. This is the basis for what philosopher Christopher Cherniak refers to as the finitary predicament confronting cognitive agency [6]. Given the vastness of information that is available for consideration at any particular moment, insofar as cognitive agency is constrained in terms of both time and

cognitive resources, it is fundamentally impossible to act functionally in the world without also reducing the world's inherent complexity in an effective manner. The solution to the problem of functional action, therefore, is achieved by means of "framing", which refers to the process of selecting only a subset of the available information based on its relevance for the situation at hand [37, 38]. Framing, in other words, is a necessary condition for the possibility of functional action in the world, and is a thus a primary function of cognition [37, 38].

Framing, however, is an inherently motivated act and not just a cold calculation [12, 13, 38], since it occurs in the service of affording functional action. Functional action is by definition goal-directed and is the means by which real-world problems are solved. It involves the careful coordination of sensorimotor activity in relation to an environment, as well as one's self and others [13]. The mechanisms by which sensorimotor activity is regulated are essentially affective in kind, involving various forms of emotional feedback, as well as feelings and moods [13], all of which are aimed at evaluating the relative potency of one's sensorimotor coordination in relation to a task for achieving contextually-relevant goals. Accordingly, behaviors (sensorimotor acts) which result in reward or which cue potential reward are experienced positively (e.g., joy, happiness, hope), and are therefore reinforced; whereas behaviors which result in punishment or which signal potential punishment (threat) are experienced negatively (e.g., fear, anger, anxiety), and are therefore extinguished [11]. Successful framing thus implies that all relevant sensorimotor affordances and affective cues are sufficiently known in the problem-solving domain at hand; functional action is afforded; and the world is thereby experienced as a place of determinate meanings, or as "predictable". As a consequence of the determinacy of the cognitive agent's frame, as well as its effectiveness in affording functional action toward the agent's goals, the agent's attendant affective state is characterized not only as relatively positive in valence, but also as relatively secure and low in anxiety. Positive affect and low levels of anxiety are, in other words, an indication that one knows what one wants, how to attain it, and also that what one knows to be sufficient for attaining what one wants is in fact sufficient.

Whereas the cognitive agent's frames of the world are static, the world as such is entropic [11, 30]. A fundamental limitation of the framing process, therefore, is that obsolescence is both necessary and inevitable. The functional utility of framing erodes whenever the agent is confronted with a wholly novel experience—an "anomaly", in a manner of speaking. During such instances of anomaly, the world's inherent complexity emerges and overwhelms the agent's cognitive structures and attendent capacity to act functionally [11, 30]. The world, in other words, becomes a fundamentally unpredictable place and its meanings are rendered obscure and indeterminate. The agent's affective state becomes characterized by relatively high levels of anxiety and emotional ambivalence (awe and terror, hope and anxiety), insofar as knowledge of how to act functionally (how to attain goals) has become confused and is no longer sufficiently predictive of either reward or punishment. The breakdown of framing—(i.e., "misframing")—is therefore a constitutive (and affectively felt) problem for the agent insofar as it necessarily renders the agent incapable of functional action. Negative affect and high levels of anxiety act to indicate that one does not necessarily know what one wants or how to attain it, while also knowing that what one has known to be sufficient for attaining one's goals is no longer sufficient.

Clearly, then, the complexity of the world must be kept at bay, insofar as functional action is imperative and depends on the sufficiency of the agent's framing for predicting the world. Confrontation with anomaly thus implies the need for amending erroneous frames by voluntarily attending to the emergent problem (anomaly) so as to facilitate a significant restructuring of the agent's overall framing. When done successfully, the agent's framing of the world regains its sense of determinacy, whereby the world's emergent complexity, together with its attendant set of anxieties and negative affect(s), is at once reduced, and functional action is afforded anew. It is but a matter of time, however, until anomaly emerges once again and the agent is forced to undergo another restructuring process of his or her framing. The circle of interpretation must turn ever onwards.

We believe that the line of argument which we have proposed describes a necessary (existential) structure of cognition, namely the ongoing hermeneutic (interpretive) circulation between framing, misframing, and reframing. We recognize how bold this claim might appear, but we believe it to be both theoretically sound [11, 13, 14, 30] and experientially compelling, insofar as (1) misframing is cognitively unavoidable and (2) cognitive reframing is imperative for regaining functionality. Furthermore, our depiction of cognition is inherently enactivist in its grounding, insofar cognition has been construed as essentially motivated (goal-directed), affective (evaluative), and procedural (sensorimotor). We follow suit with the enactivists in claiming that cognition is not a brainbound process that is functionally divorced from other bodily processes, but a form of motivated and affectively-imbued activity that is fundamentally distributed across the brain-body-environment dynamical system [13]. Cognition, in other words, is not something that one has, but something that one does in relation to a world of meaningful activity. Cognition is enacted qua the agent's embodied interaction with the world and is always bound up in the agent's practical context.

It is now possible to theorize about perception in enactivist terms on the basis of the established argument. If cognition is about functional action in relation to a meaningful world, then perception mediates cognition insofar as it constitutes the primary means by which the world is even disclosed into conscious awareness. Enactivists theorize that perceptual experience is achieved as a function of mastery of "sensorimotor contingencies" [13, 35]. Sensorimotor contingencies refer to the invariant sensorimotor structures that emerge as a function of how patterns of sensory flow, inherent in and contingent upon each distinct sense-modality (e.g., visual, auditory, tactile, etc.), covary with patterns of motor activity in relation to the attainment of a given practical aim (e.g., satiation of hunger, quenching of thirst, escaping a predatory attack, etc.) [13]. Unless the agent appropriates or incorporates into his or her procedural repertoire these invariant structures, (i.e., grows accustomed to and thereby learns how to "predict" changes in sensorimotor flow in a relevant manner), navigating the world in relation to practical aims is rendered an impossible task. Conversely, unless sensorimotor activity is motivated, and therefore bound up by practical aims, then the agent does not readily learn to perceive "objects", since the possibility of perceiving objects is necessarily tied with any object's potential relevance for satisfying practical goals (i.e., how predictive a given set of sensorimotor patterns is of either reward or punishment). Thus, perception is achieved as a matter of increased procedural familiarity with or mastery of sensorimotor contingencies and their relevance to real-world goal-

attainment. Perception is therefore a practical skill; it is not a matter of seeing the world "as it is", but of learning to see what is relevant in the world for acting functionally (attaining goals) in it [25]. In subserving the agent's cognitive aims in this way, perception is therefore also bound up (framed) by the agent's practical context in the same way cognition is. This implies, quite straightforwardly, that perception is essentially both motivated and affective. Perception is therefore a form of affectively-imbued, motivated action for mediating cognition.

3.2   Bridging the First Two Gaps

The proposed theory of perception as motivated action is firmly rooted in the enactive approach to cognitive science [12–15, 35]. Our argument suggests that perception is a matter of skillful sensorimotor coordination that is regulated by affective means, and framed according to the practical aims of the agent's problem-solving context. In this regard, the proposed theory of perception as motivated action implies an inherently pragmatic, but also phenomenologically informed treatment of perceptual reality, which runs counter to that of naïve realism. The fundamental notion underlying such a treatment is that *we do not perceive objects and then infer their meaning, but we perceive meaning and only then infer objects on this basis* [11, 30]. Recall that the first theoretical gap required an explanation as to why the perceptual (re)calibration during the onboarding process develops gradually like a skill, rather than instantaneously and mechanically as a naïve realist theory would suppose. According to enactivism, perception is achieved as a function of progressive mastery over sensorimotor contingencies through active, exploratory, and pragmatically bound participation in a world. The first theoretical gap is thereby bridged by means of our proposed theory insofar as perception is predicated on exploratory behavior and is achieved not as a matter of seeing the (virtual) world "as it is", but as a function of *learning to see what is relevant in the (virtual) world for the purpose of acting functionally in it*. Therefore, the onboarding process entails a gradual and progressive development, rather than an instantaneous one, therefore, since it demands from the individual to skillfully compensate for the discrepancy between the sensorimotor skill set that is brought into the virtual experience at its onset, predicated as it is on real-world physics, and the sensorimotor skill set that is demanded by the virtual experience, the physics of which necessarily deviate from real-world physics as a consequence of technological and engineering constraints (e.g., programming errors), on the one hand, and the fact that perception is now being mediated not just by one's body, but by an additional (augmented) physical interface (e.g., VR headset and controllers), on the other hand. The sensorimotor contingencies proper to a given virtual experience emerge as a function of the patterns of the agent's perceptual interaction with the VR content as mediated by the (augmented) physical interface, whereby mastery of said contingencies is accordingly to be achieved as a function of sustained, motivated interaction with the VR content qua its interface.

The second theoretical gap is also bridged insofar as the proposed theory advances a non-reductive synthesis of objective and subjective factors in its depiction of perceptual reality. Furthermore, it does so in a way that is both scientifically rigorous (as per its grounding in enactivism, as well as psychological science) and experientially

robust (as per its phenomenological sensitivity). In particular, the world as it is perceived, according to the argument thus far, is more appropriately regarded as a forum for meaningful action, than a collection of material objects [11, 30]. It thus follows that perception of a virtual world is therefore also the perception of a world of possibilities for functional action, in which subjective factors such as goals, motivations, and affect, play an essential part in framing what is relevant. On this basis, we find it reasonable to claim that any set of guidelines for VR design that is predicated on the naïve realist view that perceptual reality is reducible to a collection of objects in space misses the mark altogether, and cannot be said to explain or entail optimal (or meaningful) virtual experience. This is because naïve realism ultimately overlooks the pragmatic nature of perception almost entirely.

Now that the first two theoretical gaps have been bridged, what remains is a bridging of the third gap. In the following section, we utilize our enactive theory of perception as motivated action to leverage an account of flow as a state of optimal perceptual engagement with the world, within which we then ground the immersion construct and circumvent its attendant conceptual difficulties. The methodological implications of our theoretical contributions in Sects. 3 and 4 for VR research and design are explored fully in Sect. 5, all of which is done in reference to a design probe that was launched and conducted with real participants in a mixed reality (MR) setting at Carnegie Mellon University [18].

4   Redefining Immersion

Our argument in Sect. 2 was predicated on a critique of the physicalist assumptions inherent in VR, suggesting a total of three theoretical gaps in the literature which, we claimed, cannot be bridged by physicalist means alone. In Sect. 3, we articulated an enactive theory of perception as motivated action, which, we showed, is able to bridge the first and second theoretical gaps. In particular, the first theoretical gap is bridged by explaining the developmental quality of perception that is entailed in the onboarding process as a function of attaining progressive mastery over the sensorimotor contingencies inherent in the VR experience. Accordingly, the second theoretical gap is bridged by conceptualizing perceptual reality not as being comprised of a collection of physical objects, but as constituting a forum for meaningful (pragmatically motivated and affectively-imbued) action.

In this section, we aim to demonstrate how the theoretical framework we laid out in Sect. 3 can be used for also bridging the third theoretical gap, which requires an explanation of flow (and immersion) in processual terms. We begin with a discussion of the problem: In order to circumvent the conceptual difficulties associated with the immersion construct, immersion ought to be conceived of in terms of flow; however, the naïve realist assumptions inherent in VR cannot account for flow in processual terms and must therefore be replaced if a non-problematic (flow-based) definition of immersion is to be recovered. We then turn to recent work in cognitive science on the flow state which conceptualizes flow as a process, that is, as an "insight-cascade" [17] whereby the hermeneutic circle comprised of framing, misframing, and reframing is, in a manner of speaking, "ramped up". Next, we integrate this cognitive scientific account

of flow into our enactive framework and argue that flow is a marker of optimal perceptual engagement with the world, that is, flow implies a temporary enhancement of the very processes of mastery-attainment over relevant sensorimotor contingencies. We conclude by grounding the immersion construct in our enactive conception of flow, thereby showing how such a grounding helps to circumvent the various conceptual difficulties associated with immersion. In Sect. 5, we proceed to articulate four methodological implications of our theoretical contribution for VR design.

4.1   Formulating the Third Theoretical Gap

The golden standard of VR currently is determined by the degree to which a virtual experience can induce a state of presence within a participant, which is broadly defined as the subjective feeling or illusion of being in one place, when one is in fact physically located in another. It is generally stipulated that presence is achieved partly as a function of the objective immersive properties of the VR hardware, such as the number of built-in sensors, and partly as a function of the quality of sensory stimulation provided by said sensors (i.e., the degree of fidelity preserved between the physics of the VR and those of the real world). The design goal of VR, in other words, aims at the maximization of presence through the simulation of physically realistic experiences as mediated by the objectively immersive VR hardware. At face-value, a presence-predicated design rationale such as this one might not seem very objectionable. However, upon closer philosophical scrutiny, two challenges are revealed—the first, a methodological challenge, and the second, a conceptual challenge.

First, the assumption that presence is perceived as a matter of course if only the VR hardware is sufficiently objectively immersive and the physics simulated in its attendant sensory stimuli are sufficiently (physically) realistic, is fundamentally predicated on a naïve realist theory of perception. As we have already argued in Sect. 3, however, perception is not a passive act, but rather a matter of skillful mastery over sensorimotor contingencies through exploratory interaction with a meaningful world. The onboarding process exemplifies the developmental and skillful character of perception in VR and therefore constitutes a counterfactual to the naïve realist assumption that the perception of presence can be causally reduced to the conjunction of the objective immersive properties of the VR hardware, on the one hand, and the quality of sensory stimulation thereof, on the other hand. Be that as it may, though, a more fundamental issue with treating presence as the golden standard of VR is arguably that it is object-centric (Newtonian), rather than action-predicated (pragmatic), whereas human perceptual reality, as we have argued in Sect. 3, is more appropriately to be understood as a forum for meaningful action than a collection of objects. In other words, making the design goal of VR the maximization of presence seems to neglect the pragmatic dimension of perceptual reality. It becomes unclear, then, how such a phenomenologically impoverished design ideal can be used for explaining, or even generating, optimal (and meaningful) virtual experience.

Second, as mentioned in Sect. 2, the construct of immersion is defined not only as an objective property of the VR hardware, but alternatively as the subjective state or feeling of "being caught up in another world". The distinction between subjective and objective immersion has consequently led to various confusions as to its use within the literature.

For example, it has been difficult to discriminate between immersion in its subjective sense and presence due to their apparent overlap in meaning on the definitional level [8]. Additionally, although the distinction between objective and subjective immersion seems to be a conceptually useful move, it nevertheless raises difficult questions as to the relationship between immersion (in both senses), on the one hand, and presence, on the other hand. The fact that objective immersion is not predictive of subjective immersion (as was argued in Sect. 2) further complicates the conceptual boundaries of the constructs at hand. As a consequence of such conceptual fuzziness, problems emerge with attempting to clearly operationalize these constructs in an experimentally useful manner. Subsequently, causal modelling of the relations between presence and immersion (in both senses) is rendered an incredibly difficult task, which cannot be resolved through statistical means alone, since correlational data can neither imply causation, nor determine the direction of causation between related variables.

In response to this conceptual challenge, we follow suit with Mütterlein in advocating a flow-based approach to subjective immersion [8]. Specifically, we define immersion as the "subjective experience of feeling totally involved in and absorbed by the activities conducted in a [virtual] place or environment", which thereby makes it possible to clearly distinguish the pragmatic, task-related elements of VR experience from the objective elements with which presence is concerned (e.g., the surrounding space and objects, social agents, and self). In other words, we believe that a conception of immersion-as-flow should sufficiently address the demarcation issue with presence and subjective immersion: Presence becomes concerned with the subjective experience of "realness" as it pertains to virtual *objects* (e.g., items, bodies, spaces, etc.), whereas immersion becomes concerned with the subjective experience of "meaningfulness" as it pertains to virtual *actions* (e.g., tool-use, problem-solving, navigation of a map). The renewed sense of clarity brought about by this flow-based conceptualization of immersion should then afford more experimentally robust operationalizations of objective immersion, subjective immersion, and presence, all of which can consequently be subjected to more rigorous causal analyses.

Accordingly, the adoption of a flow-based approach to subjective immersion not only helps to address the various conceptual difficulties currently besetting the literature, but also helps to motivate the methodological shift we aim to make from an object-centric approach to an action-predicated approach to VR. For provided that a naïve realist theory of perception in principle cannot account for flow—since flow is fundamentally a process, whereas naïve realism (given its Newtonian heritage) conceives of the contents of perception as material objects—a conception of immersion-as-flow is not only conceptually appropriate to pursue, but methodologically necessary. In what follows, we summarize contemporary work on the cognitive science of flow in preparation for advancing a processual, action-predicated (pragmatic) alternative to the current object-centric (Newtonian) approach to VR research and design.

4.2   The Cognitive Science of Optimal Experience

The psychological literature explains the flow state as a consequence of a tight coupling that is obtained between an agent and his or her environment. This tight coupling is mediated by clear and contiguous environmental feedback in response to the agent's

performance on a task, whereby errors are highly diagnostic of whether and to what degree task demands are being met, as well as what "adjustments [are] needed in order to maintain performance" ([17], p. 311). Importantly, it must also be the case that the task is perceived by the agent as challenging, but that its level of difficulty exceeds the agent's skill level only by a small margin. If the task becomes too difficult, then the agent is overwhelmed by feelings of anxiety and frustration such that the coupling of agent and environment that is necessary for attaining flow is lost ([17], p. 311). Conversely, if the task becomes too easy, then the agent experiences boredom and thus loses the motivation to sustain engagement—as a consequence of which, the tight coupling of agent and environment is lost once again. Flow is therefore described as the state of "optimal" experience because, when inhabiting it, the agent's skill level is continually being pushed to its limits, resulting in the emergence of a "skill-stretching" function ([17], p. 311). Vervaeke et al. [17] describe "skill-stretching" as "a system of learning where the process of meeting and overcoming one challenge breeds a new and more developed skill set, in turn affording the ability to take on a still more difficult set of demands" (p. 311). As such, the self-perpetuating sense of motivation that is often felt during flow is experienced from being consistently challenged while nonetheless reliably overcoming such challenges along the way [5].

Whereas psychological accounts have typically described flow at the level of interaction between agent and environment, Vervaeke et al. [17] offer a description of flow at the level of cognitive processing, which makes their depiction of flow particularly amenable to theorization in enactive terms. We therefore turn to an explication of Vervaeke et al.'s cognitive scientific account of flow, which we aim to integrate with our own theory of perception as motivated action. As we will see, pursuing such a synthesis will provide us with the grounding that is needed both in order to address the conceptual difficulties associated with immersion, as well as to motivate the much needed methodological shift toward an action-predicated (pragmatic), rather than object-centric (Newtonian), approach to VR.

Vervaeke et al. [17] conceptualize the cognitive basis of flow as an insight-cascade, a concept which requires some unpacking, starting with the notion of "insight". The realization of an insight is often described as an "aha!" moment and is also commonly depicted with the metaphor of a lightbulb going on inside one's head. Importantly, insight learning constitutes a qualitatively distinct mode of learning: Whereas in conditional learning (e.g., classical or operant), the rate of learning is often incremental, in insight learning, it is characterized by the abrupt or spontaneous realization of a solution that is preceded by a period of (prolonged) impasse [17]. During insight problems, the solution is not achieved by straightforward means, such as by recalling relevant content from memory. Indeed, what typically characterizes an insight problem is precisely that one's prior knowledge interferes with the realization of a relevant solution to the problem [14]. In other words, when faced with an insight problem, the agent experiences a fixation in how the problem has been (incorrectly) framed and must realize a solution only by overcoming the impasse caused by the misframing. The insight is thus realized as a consequence of a shift in the agent's attentional processing of the situation, which occurs at the level of procedural (not propositional) processing, causing the agent to break out of the erroneous frame and to form a novel framing of the situation by which a solution is finally afforded [14, 17].

Given this depiction of insight problem solving, Vervaeke et al. [17] argue that the flow state emerges as a dynamical system whereby the act of solving one insight problem immediately gives way for and creates an additional insight problem that can be solved, a process which is then iterated and sustained over a span of time—granted that the proper learning conditions are in place (i.e., tight environmental feedback and optimal task difficulty). A cascade of insightful processing emerges as a consequence, causing a "stretched out 'aha!' moment". The ongoing realization of insight is what affords the "skill-stretching" function of flow. In it, the agent's competence at tackling problems of the kind being faced is continually being improved and stretched beyond its limits with every successive instance of insight. This is one of the reasons why flow is also referred to as a state of optimal *learning* or *engagement* [17].

Skill-stretching is not the only core characteristic of the flow state, however. Recall that during flow, the agent experiences a sense of ineffability (non-deliberateness) throughout the engagement process, as well as a paradoxical sense of immersion in the task, whereby the quality of engagement is both effortful yet effortless. Vervaeke et al.'s [17] cognitive account of flow explains immersion in terms of fluency, and ineffability in terms of intuition. Fluency refers to the sense of ease or difficulty associated with a cognitive process. In citing the work of Topolinski and Reber [24], Vervaeke et al. [17] suggest that the subjective sense of fluency accompanying a cognitive process might in fact be correlated to "the actual degree of ease of processing occurring at the neural level" (p. 312). Thus, whereas specific instances of insight problem-solving are accompanied with discrete moments of enhanced feelings of fluency, they argue, "it follows that a cascade of insights would naturally yield an accompanying and ongoing stream of positive subjective affect, reinforcing a sense of meaning in one's processing—flow phenomenologically equates to an experience of extended fluency" (p. 312). The paradoxical sense of immersion associated with flow is thereby explained as a function of extended fluency, wherein individual moments of frustration caused by impasse (low fluency) are spontaneously interrupted by moments of insight (high fluency), which yield satisfaction, forming, as a consequence, a positive feedback loop of self-perpetuating, self-motivating, and reliable engagement with the task at hand.

Vervaeke et al. [17] accordingly explain the ineffability of the flow state by grounding flow in intuition. In reference to work by Hogarth [23], they describe intuition as a product of implicit learning—which is "*tacit*, as opposed to *deliberate*"—and as "effortless, reactive, and producing 'approximate' responses" (p. 321). Provided that implicit learning and flow are both non-deliberative and ineffable/procedural (rather than voluntary or propositional) processes, grounding flow in implicit learning seems like an appropriate conceptual strategy. The obvious challenge, though, is that whereas flow is a state of *optimal* experience, implicit learning as such appears to be a *suboptimal* process on the whole. For, on the one hand, implicit learning suffers from the problem of "over-fitting", which occurs when "correlational noise from the environment is interpreted as being causally relevant to the pattern of action" (p. 321). During flow, however, over-fitting is not a problem, since, if it were, then it would fundamentally disrupt the insight-cascade, thereby rendering the flow state a practical impossibility. On the other hand, though, implicit learning is confined primarily to tracking *actual* patterns in the environment, whereas flow involves the adaptive

tracking of and selection from *possible* patterns for dynamically affording functional action (pp. 321–322). Vervaeke et al. [17] insightfully note that the conditions for acquiring sound intuitions, namely, separating "causal signal" from "correlational noise", happen to mirror those for cultivating a state of flow: "A system of learning that tightly couples actions and environment with timely feedback—thus providing high error diagnosticity—is a system conducive to cultivating flow and good intuitions" (p. 322). On this basis, the authors advance a conceptual synthesis of flow and implicit learning in which they propose that flow "is optimal for implicitly learning complex patterns in the environment and distinguishing them from correlational ones while exploring possibilities of action and learning" (p. 322). Such a conceptual synthesis helps to explain the non-deliberative quality of the flow state without reducing flow to a set of imprecise, automatic processes concerned primarily with tracking patterns in actuality. The authors summarize this point eloquently: "Flow is a system of processing and cultivating causal pattern recognition in which cognition is stimulated to explore possibilities of action. These two elements are interdependent: exploring possibilities allows one to distinguish between actual causation and mere empirical generalization. In turn, zeroing-in on causation helps guide the insight away from being illusory or fantastical" ([17], p. 322).

The cognitive scientific account of the flow state reviewed here explains the core features of flow in cognitive terms. Skill-stretching is a qualitatively distinct mode of learning that emerges as a function of the insight-cascade, in which insight problems, on the one hand, and insight problem-solving, on the other hand, enable one another in a mutually affording fashion, sustained over a span of time. The immersive process, and its attendant phenomenology, are accordingly explained as a function of sustained fluency. Finally, the ineffable character of flow is explained by grounding flow in intuition, a non-deliberative cognitive process, whereby flow is conceived of as a procedurally-driven, optimal form of implicit learning.

4.3   Enaction, Perception, and Flow

In building up to a flow-based reconceptualization of subjective immersion, we must ensure that our argument retains a level of philosophical consistency throughout. In order to ensure this, we must therefore ground the cognitive account of flow outlined in Sect. 4.2 within our own theory of perception as motivated action. Demonstrating such a grounding should not only guarantee a necessary degree of coherence, but should also help to motivate the methodological shift we are attempting to make in this paper toward an action-predicated, pragmatic approach to VR.

The cognitive account of flow as an insight-cascade is readily interpretable through our enactive lens. In Sect. 3, we argued that, as per the finitary predicament, the hermeneutic circulation of framing, misframing, and reframing is a necessary existential structure of cognition. Taken in these terms, an act of insight becomes understood as an instance of spontaneously reframing a problem frame, which thereby affords functional action and enables the agent to regain fit with the environment after a period of frustration and impasse. It follows that, if flow is indeed a cascade of insights whereby learning is optimized and one's skills are continually stretched beyond their limits, then the flow state constitutes a "hermeneutic hypercycle" whereby the necessary circulation

of framing, misframing, and reframing, is, in a manner of speaking, ramped up and sustained processually over a period of time. The cultivation of sound intuition by way of flow thus translates into the cultivation of mastery over sensorimotor contingencies, wherein not only is greater perceptual mastery obtained over one's engagement with the environment, but the very processes of mastery attainment are themselves temporarily deepened and enhanced. Flow is, in other words, the instantiation of an optimal form of perceptual engagement with the world, of which "skill-stretching" is an emergent function.

Our theory of perception also describes the attendant phenomenology of the flow state as being constitutive of its own class of meaning, which is distinct from how the world is experienced during instances of accurate framing (whereby functional action is afforded), on the one hand, and misframing (whereby functional action is not afforded), on the other hand [30]. Specifically, flow is a state of engagement in which the world as a forum for action is in the active process of being transformed from a place of indeterminate meaning, wherein functional action is impeded, to a place of determinate meaning, wherein functional action is afforded [30]. The ineffability of flow, accordingly, is accounted for by the fact that flow is fundamentally a procedural (non-propositional, non-deliberative) process constituted by a chained sequence of sensorimotor breakthroughs with respect to a meaningful environment. The immersive tendency of flow, subsequently, is explained as a function of an optimal degree of indeterminacy—and experienced anxiety—that characterizes an agent's perception of the world, as well as both the rate and clarity of sensorimotor feedback by which a sense of determinacy—and an accompanying feeling of security and confidence—is salvaged from one's interaction with the indeterminate. As the literature on flow clearly states, the difficulty of the task must be just beyond one's own skill level in order for the flow state to obtain. In other words, the indeterminacy of the world and its attendant anxiety must remain at an optimal level throughout, so as to simultaneously beckon the agent's meaningful engagement without actually overwhelming his or her capacity to engage meaningfully.

Having in this way described the flow state's three main attributes (skill-stretching, ineffability, and immersion), we can now claim that the cognitive account of flow has been sufficiently grounded in our proposed theory of perception as motivated action, and is, as a result, conceptualized in enactive terms. A level of philosophical consistency has therefore been ensured, insofar as the various concepts used throughout our discussion (e.g., perception, action, affect, cognition, flow) have all been grounded in an enactive framework. We can finally proceed to the last step of our argument, where we conceptualize subjective immersion in terms of flow.

4.4   Immersion-as-Flow: Toward an Action-Predicated VR

Throughout this article, we have taken deliberate steps to build a disciplined critique of the physicalist presuppositions implicit in VR. The current design goal of VR is to maximize presence by simulating physically realistic experiences that are mediated through objectively immersive hardware. We have challenged the validity of this design goal on two distinct, yet interrelated fronts. First, such a design goal tacitly subscribes to a naïve realist theory of perception, which fundamentally cannot account for the gradual and skillful character of perception during onboarding in VR. Second, it

commits to a Newtonian view of reality as a collection of physical objects, and it assumes, as a consequence of this commitment, that the objects of perception are reducible to the objects of the (Newtonian) physical world. Since the pragmatic structure of perceptual experience is, in principle, precluded by such an object-centric approach to VR design, we therefore proposed that a methodological shift be made toward an action-predicated approach instead. Another part of the "third theoretical gap" which we described pertains to conceptual issues with adequately delimiting the core constructs used in VR research and design: Presence, objective immersion, and subjective immersion. We are now ready to advance the final step of our argument, which aims to bridge the third theoretical gap. Specifically, we ground subjective immersion in flow and demonstrate how doing so simultaneously addresses the conceptual difficulties posed by the third gap, as well as how it engenders the much needed methodological shift toward an action-predicated approach to VR.

If subjective immersion is to be conceptualized in terms of flow, it follows that immersion-as-flow is fundamentally a process of cultivating sound intuition in relation to virtual tasks. Accordingly, immersion-as-flow is attained in VR when (1) environmental feedback in the VR is clear and contiguous with one's patterns of engagement, and is therefore highly diagnostic of one's performance; and (2) the associated difficulty of the virtual task remains optimal throughout, that is, it stays just beyond one's own level of skill, thereby affording an emergence of "skill-stretching". It should be evident that instantiating (1) and (2) in a VR setting is not a matter of simulating a physically realistic environment *per se*, contrary to what the design goal of maximizing presence would prescribe. Rather, it is a matter of designing an experience wherein functional action can be clearly and reliably afforded from a user's point of view. Specifically, in treating the virtual environment as a forum for action, in order to facilitate or even maximize immersion-as-flow, VR creators must (i) define the possibilities for action (sensorimotor coordination by way of the VR interface) in relation to the virtual environment, (ii) clearly demarcate those possibilities which count as rewarding from those which do not, and (iii) ensure that conditions (1) and (2) are in place while the user is in the process of learning what the relevant (sensorimotor) possibilities are and how they can be enacted. Normally, (iii) is realized through clear, guided instruction, presented in the form of a tutorial (explicit or implicit) during the onboarding process.

If successful, the process of immersion should result in mastery over the sensorimotor contingencies inherent in the VR experience. As a result of progressive mastery, the VR interface becomes increasingly incorporated into one's perceptual skill set and the act of perceiving thus becomes increasingly intuitive and "immediate" in the way it feels for the user. With sustained immersion, in other words, the "headset-as-blindfold" eventually (and rather unnoticeably) becomes a "headset-as-window-into-the-virtual-world" (recall from Subsect. 2.1). A rather curious implication of our framework is that the cultivation of "sound intuitions" via sensorimotor mastery in this way constitutes the cognitive basis for the experience of presence in VR. The presence of various objects in the VR experience is therefore to be understood as a function of the accumulation (and incorporation) of "sound intuitions" regarding relatively stable (constant) sensorimotor patterns available in the proximal virtual environment. Presence, therefore, is a consequence of successful and sustained action-predicated immersion in VR, insofar as sustained and successful immersion does indeed yield sound intuitions.

By having adopted a flow-based approach to subjective immersion, greater operational rigor is now also afforded with regards to the core constructs of VR: Presence, objective immersion, and subjective immersion (immersion-as-flow). First, by conceptualizing subjective immersion in VR in terms of flow, it has become possible to clearly distinguish the objective elements of VR experience (e.g., items, bodies, spaces, etc.), with which measures of presence are primarily concerned, with the pragmatic, task-related elements of VR experience (e.g., tool-use, problem-solving, spatial navigation, communication), with which measures of immersion-as-flow are primarily concerned. Furthermore, the fact that objective immersion is not predictive of subjective immersion (recall Subsect. 2.1) can now be explained as a consequence of the fact that what matters for immersion-as-flow is not the degree to which one is *objectively* immersed, but rather the quality of information that is communicated through the immersive medium *and* whether and to what degree such information is conducive of flow. In this way, the proposed account of immersion-as-flow circumvents the conceptual issues implicated by the third gap.

In addition, though, the proposed account of immersion-as-flow is grounded in a theory of perception whose presuppositions are neither Newtonian, nor naïve realist in kind. In this way, the methodological challenge posed by the third theoretical gap is similarly circumvented, and in a rather straightforward manner: Perceptual reality in VR is enacted through embodied, practical, and exploratory engagement with a virtual world, whereby one does not come to perceive the virtual world "as it is", but rather learns to perceive what is relevant for attaining one's practical purposes in it. As an alternative to the current design goal of VR (i.e., maximizing presence), which is fundamentally object-centric, we thus propose the action-predicated goal of maximizing subjective immersion. Our claim is not that the experience of presence is unimportant for VR, but rather that its importance is only secondary to that of maximizing subjective immersion (i.e., virtual flow). As such, we believe that the golden standard by which to measure the quality or "success" of VR is the degree to which the virtual experience can be said to be conducive of subjective immersion in the user. Needless to state, such an approach makes intuitive sense on yet another level. For insofar as flow constitutes a criterion of optimal experience, it follows, therefore, that immersion-as-flow may thereby constitute a criterion of optimal virtual experience.

Our theoretical argument is now complete. We began with a critique of the current paradigm within VR, which, we argued, is tacitly physicalist in its grounding. Next, we claimed that because of its physicalist assumptions, VR research is beset by three theoretical gaps which cannot be bridged by physicalist means alone. We proposed an alternative, action-predicated (pragmatic) approach by drawing from enactive cognitive science, which we claimed could help to bridge the three gaps and circumvent their attendant challenges. In our estimation, all the gaps have now been bridged, and their challenges, circumvented. With our theoretical contribution now realized, we must demonstrate its practical utility by illustrating the methodological implications for VR design which from it follow. In the following section, we articulate our implications, and, in Sect. 6, we briefly summarize our findings and conclude by highlighting potential avenues for future research.

5   Methodological Implications for VR Design

In this section, we demonstrate the practical utility of our theoretical contribution by proposing a set of methodological guidelines for VR creators. We identify four essential elements of VR experience—(1) Onboarding, (2) Immersion, (3) "Offboarding", and (4) "Experiential optimization"—and organize our methodological commentary into four subsections, each of which addresses one of these elements. Rather than dictating *what* to design, though, our methodological merits are meant instead to model a *general approach to the very process of designing* head-mounted, immersive experiences. To this end, we begin first by examining a design probe, *Wake*, that exemplifies a real-world implementation of our proposed approach. Our subsequent discussion of *Wake* is then complemented with and grounded in the theoretical narrative laid out across Sects. 3 and 4.

5.1   Design Probe: *Wake*

Wake is a facilitated mixed reality (MR) experience, created in 2018 by Anna Henson in collaboration with the Pittsburgh-based multidisciplinary performance duo, *slow-danger* (Anna Thompson and Taylor Knight), with research assistance from Qianye (Renee) Mei and Char Stiles. *Wake* is a site-specific, participatory, movement-based installation facilitated by two dancers, in which one participant in-headset (using the HTC Vive Pro room-scale virtual reality system, Vive spatial trackers, and an Intel RealSense depth camera) navigates a walkable virtual environment, which corresponds in size and layout with the physical environment. The participant interacts with both virtual and physical (tangible) objects, and a co-present dancer, who is tracked and rendered photographically in real time in the headset using the head-mounted depth camera (Intel RealSense). The participant is initiated into the experience by one un-tracked dancer who serves as a facilitator (managing the hardware, providing instructions), and later encounters and interacts with the second dancer through visual gestures, physical touch, and verbal dialogue. *Wake* engages concepts of embodied interaction [21] and social presence within hybrid mediated environments.

A user study (N = 25) was conducted to investigate participants' cognitive and affective experience during *Wake*. Qualitative data were collected through semi-structured phenomenological interviews and a standardized self-report for emotional states [22], and quantitative data were gathered through spatial trackers between the participant and co-present dancer, which was analyzed using proxemics. The chosen participants constituted a sample of convenience, as the aim was to explore general principles rather than to experimentally study the effects of specific variables. Participants ranged in age from 21–48 (mean age = 28.5), were recruited mostly from universities in the Pittsburgh area, and consisted of both VR novices and VR developers. All participants had a baseline of using technology to communicate with others, as co-located (located together in the same physical and virtual space), co-presence, and non-verbal communication within a hybrid immersive media environment were significant areas of investigation in this design probe. *Wake* incorporated a user study to explore and interrogate methods of embodied interaction for VR, and to dialogue with participants about

concepts developed through Anna Henson and *slowdanger's* collaborative, practice-based research process.

5.2   Onboarding and Pre-immersion

The Onboarding Process. In every head-mounted virtual reality experience, a threshold must be crossed from perceiving the world without wearing a headset, to putting the headset on and engaging in virtual content. This process, onboarding, is crucial for the participant to become fully immersed in the virtual experience. The hardware itself is an inescapable physical reality, though, and in the case of room-scale VR systems (HTC Vive, Oculus Rift, Sony Playstation VR), it is reasonably bulky. The headset and any other worn sensors are in intimate relationship to the participant's body to allow for engagement in the virtual content. The hardware's form factor will continue to decrease in physical size as technology evolves, but, presently, the hardware completely covers the eyes and a significant portion of the face of the wearer. This is the "headset-as-blindfold" phenomenon articulated in Sect. 2. These hardware factors can trigger discomfort, disorientation, or other negative responses in participants across physical, affective, and social levels, if not attended to properly during onboarding. If the headset and other hardware are not appropriately incorporated into the participant's skill set, and consequential affective concerns are not addressed from the outset, this can render moot the actual content in the virtual experience. If the onboarding process (regarding hardware worn on the body, virtual interface, and, if relevant, relationship to other people in the experience) is confusing, nonconsensual, or abrasive, the participant can become distracted or may entirely disengage from the experience right away.

Smoothing the transition from outside to inside the headset is thus a primary concern of the early stages of a VR experience. The experience design should attempt to counteract the discomfort or distraction of the hardware, to help foster a sense of safety or trust, and to cultivate intuition in the participant's ability to perform physical movement wearing the hardware, which is necessary for the subjective feeling of immersion.

Mastery and Scaffolding. VR is a medium with high cognitive load. The early moments of perception and interaction (with both the hardware and the content) in a VR experience are crucial to cultivating a sufficient degree of intuition so that the participant can become safely and fully immersed. The terms *scaffolding*, *affordances*, *mastery*, and *discovery* are all germane for conceptualizing the design of a participatory experience, virtual or otherwise, and can be used for understanding how intuition is cultivated during onboarding, as well as how immersion is made possible as a result.

Scaffolding is used here to denote the ways in which a participant is instructed through a task, which in turn makes the task easier and adds fluency to the learning process. *Affordances* are the possibilities for action that a participant perceives, and through appropriate scaffolding, affordances of greater relevance become available to the participant. Put differently, affordances are perceptual frames, and, so, functional action is enabled based on the perceptual affordances available at the time of interaction. *Mastery* denotes competency with a skill or action, whereby relevant affordances become progressively more intuitive. Mastery is achieved through

repetition of motivated, exploratory behavior that successfully enables functional action (i.e., yields reward and/or avoids punishment). Once a basic level of mastery is attained, the participant's engagement with the task becomes more intuitive and thus attains greater processing (and experiential) fluency. Subsequently, greater immersion is achieved and the realization of a flow state within the virtual experience becomes more probable.

During the initial mastery stage of a VR experience, the participant acquires basic perceptual (sensorimotor) skills for the VR environment. Within the *Wake* design probe, two areas of "orientation" were found to be crucial for the participant's experience: First, attending to the participant's own sense of embodiment once the hardware is worn, and, second, seamlessly coupling this hardware-affected sense of embodiment with the perception of virtual space and objects. To address these participant needs, *Wake* developed: (a) Physical Orientation Exercises (POEs) and (b) Virtual Orientation Exercises (VOEs). The following discussion will elaborate on the POEs and VOEs used in *Wake*, whereby s*caffolding* is done verbally through instruction, and is socially negotiated.

Physical and Virtual Orientation Exercises. POEs and VOEs in *Wake* are designed to help the participant gain facility with the embodied situation of wearing the headset, and also the physical and visual tools they will use later in the virtual experience. The POEs, which are conducted while the participant perceives darkness in the headset, consist of breathing, sensory awareness, and simple directed movements. When enacted, these actions help the participant to feel greater proficiency over their own bodily proprioception, increase feelings of physical safety, and help the participant to trust the facilitator, which can help enable the participant to move through the experience with more comfort and receptivity. Trust in *Wake* appears to emerge through this sort of facilitated embodiment. One participant stated, "The short breathing exercises at the beginning helped refocus my body, so I felt more comfortable wearing the headset, and the initial feeling of apprehension started to fade away" [18].

VOEs, on the other hand, are meant to acquaint the participant with the "rules" of the virtual experience. More specifically, this entails being introduced to, and later mastering, the possible affordances that are available in the virtual environment through the interface. Transitioning seamlessly from POEs to VOEs, the participant in *Wake* begins to see virtual objects (translucent white rocks) which correspond to their tracked wrist movements (wearing the Vive trackers), and a green rope between the two rocks. The participant is instructed to "play" with these virtual objects, by moving their arms and witnessing the interaction of the rope, which moves dynamically and with real-world physics. Additionally, the participant soon encounters three red spheres, which appear one at a time at a height of about 1.4 m, to which the participant is connected via the same green dynamic rope encountered earlier. These spheres respond to interaction in a similar manner.

Importantly, the real-world dynamics of the rope, and the height at which the spheres appear (generally at a level where participants can look straight ahead, not up or down), help to create intuitive physical interactions in virtual space. Many discussions of virtual interaction design articulate the great importance of dynamic, responsive movement which believably corresponds and contributes to the bodily sensations and movements

of a participant. Through sustained interaction, the participant acquires mastery over these basic virtual tools and physical movements, which provides motivation, positive feelings, and fluency with the subsequent parts of the experience. Once basic mastery over the VR experience is acquired in this way, the transition from onboarding to immersion can be said to have begun.

5.3   Immersion and Discovery

Basic mastery over relevant skills is necessary for flow, which is realized when task difficulty is just beyond an individual's skills. Having thus acquired basic mastery during the onboarding process via POEs and VOEs, the participant is now prepared for a more immersive experience. Immersion-as-flow within VR can thus be facilitated by introducing complexity into the VR scenario, thereby progressively increasing the cognitive and sensorimotor demands of the task(s) at hand. The addition of complexity might, for instance, entail introducing a novel challenge into a game or a puzzle which requires the participant to enact a creative synthesis of two, previously known problem-solving strategies, into a novel, composite strategy (by using a tool in an entirely novel manner to solve a problem). Situational complexity, though, must neither overwhelm the participant's ability to cope skillfully, nor be exceeded altogether by the participant's practical know-how. But should the demands of the task only slightly exceed the participant's level of skill, the complexity of the situation will help to garner and maintain participant interest and motivate the participant to engage in exploratory behavior, or *discovery* (i.e., the motivated discovery of possibilities for action). The design goal during the immersion stage is therefore to strike and sustain an optimal balance between the cognitive (and sensorimotor) demands of the situation and the available skills of the participant, so as to facilitate ongoing interest, engagement, and discovery of possibilities over time.

In *Wake*, this was achieved through the use of theatrical staging techniques (such as directional lighting), as salience cues, as well as verbal instructions, to direct the participant's attention so as to continually scaffold their learning throughout the course of the installation. The virtual experience began in complete darkness, with POEs as the main emphasis. VOEs were then introduced, which aimed to teach the participant how to effectively coordinate their physical movements in relation to the objects that would appear sequentially inside the virtual environment (e.g., rocks, rope). Through each subsequent stage of the experience, a novel function was introduced with the appearance of a new object or aspect of the virtual space (e.g., a path), which could only be accommodated and mastered by synthesizing previously learned behaviors (sensorimotor acts) into composite and more complex behavioral patterns. Through the progressive introduction of novel functions and possibilities for action, the virtual environment became an increasingly complex arena for the participant to act in; and through the scaffolding of the participant's learning, the participant's skills were continually stretched to match the growing demands of the situation. The process of ongoing discovery in *Wake* culminated into a moment whereby the virtual object with which the participant had already been interacting was revealed to have been under the direct, physical control of the dancer all along (e.g. a particle system controlled by the two Vive trackers worn on the dancer's wrists). Particularly, this realization occurred as

a live capture of the dancer was rendered in the virtual environment (more on this in Subsect. 5.5), overlaid on top of the virtual object. The co-located dancer was hence transformed into a virtually co-present agent, thereby changing the meaning of the participant's virtual situation, and making possible a whole new kind of interaction altogether.

5.4   Post-immersion Offboarding

An immersive VR experience does not end abruptly once the user removes the headset. Just as becoming immersed entails a transition period (i.e., onboarding) in which the participant's sensorimotor systems must calibrate to fit the sensorimotor demands of the virtual experience, the post-immersion experience likewise entails a transition period, an "offboarding" process, if you will, which entails a reorientation to the familiar. As designers, we must therefore acknowledge that the participants will go through a reorientation period after taking off the virtual reality headset and other equipment, in which they will need to process or "decompress" from the experience. This means that we should create a scenario in which such processing may occur, either a quiet place for reflection, a medium for expression (such as a guest book), or a place to talk with other participants.

In *Wake*, offboarding involved the administration of semi-structured phenomenological interviews inquiring into four general aspects of the participants' experience: (i) Bodily sensations, (ii) Emotions, (iii) Relationship to the dancer/facilitator, and (iv) Interaction and spatial design. An interesting observation that was drawn regarding participants' experience of offboarding in *Wake* was that there was a clear shift in vocal tone and language with most participants over the course of the interview. Their descriptions were initially highly intuitive and centered around bodily feelings and sensations, but become increasingly more analytical and deliberative as the interview progressed. This transition from intuitive to deliberative language suggested that the the participants' attention was initially largely preoccupied with various sensorimotor and embodied aspects of their immersive experience, and that the interview facilitated a sort of processing and integration of these aspects of their experience into their consciousness post-immersion.

5.5   Experiential Optimization

We have extensively argued that the design goal of VR should be the maximization of subjective immersion, rather than presence. We find it methodologically important here to identify and address an optimization issue, which we call "experiential optimization", pertinent to the realization of this design goal. Experiential optimization involves a trade-off relationship between objective immersion, on the one hand, and subjective immersion, on the other. More specifically, it appears that as the degree of objective immersion afforded by a given VR hardware (i.e. worn sensors, controllers, headset) is maximized, so increases the degree to which user perception in VR becomes mediated. Consequently, perceptual engagement and interaction with VR content becomes increasingly counterintuitive or clunky. In other words, there *prima facie* appears to be a limit on the degree to which objective immersion can be maximized before the design

goal of maximizing subjective immersion becomes compromised. This is not to say that objective immersion should be forsaken altogether, since it is, after all, a necessary feature of immersive VR experiences. Rather, given that both subjective immersion and objective immersion are essential for immersive VR experience, and that there is a necessary trade-off between these two kinds of immersion, the methodological principle here becomes not the maximization of one kind of immersion over the other, but rather the optimization of the trade-off between the two.

In our estimation, the way to experiential optimization cannot be prescribed in a manualized manner, but must be determined (or discovered) on a case-by-case basis, depending on what the VR experience in question is meant to express or engender. Optimizing the trade-off between objective and subjective immersion can mean including such tools and sensors as tracking devices, haptics, or artificial intelligence, as part of the design. However, due to the interference caused by added layers of mediation, part of a designer's job is to know which virtual elements to lean into, and which virtual elements to leave out of the equation in order to create an experience in which subjective immersion (immersion-as-flow) is properly facilitated or achieved. Optimization thus might even become a matter of also engaging the un-mediated sense modalities of an individual with the VR content, as a way to creatively sidestep technological limitations of the hardware and interface in favor of enriching the experience design and leading to greater immersion. We thus turn to a discussion of how experiential optimization was achieved in *Wake*, with regards to the problem of representing others in VR.

Representing Others in VR. Visually representing others in co-present VR experiences can be done in many ways, but the vast majority of these involve avatars (i.e. a human-controlled, computer generated representation of a person or character). Avatars exist, for varying purposes, on a scale of realism to abstraction, and many experiences using abstract or fantastical avatars can be said to be highly successful. However, representing a unique individual with a high level of photographic realism is currently a critical question. Recent developments in 3D modeling and scanning (i.e. photogrammetry) have made highly detailed renderings and photographically-based captures of individuals possible; yet a 3D scan is simply a static, unmoving mesh, and even the most advanced, rigged 3D model of a human still confronts the "uncanny valley", or the repulsion experienced when faced with a humanoid representation which is *almost-but-not-quite real,* or *strangely familiar* [31]. A 3D scan of a person may thus be *photographically realistic*, but unless the scan can *behave realistically*, its embodied expression will not convey "aliveness", feelings, or intent in a manner that is intuitive or realistic. Put differently, real-time, intuitive, social communication between people within immersive media is enabled when participants can see, hear, and respond to others in a believable and instantaneous manner.

Volumetric capture is a video technique that utilizes synced RGB and depth streams of a person or environment to render in 3D (i.e., as a mesh or point cloud) its subject. The captured material can be edited and used in VR experiences in a similar way to traditionally filmed content. With high resolution capture, subtle facial expressions and body language are made visible and therefore available to the participant, in a similar way to our real-world interactions in the physical world. This

technique thus solves the problem of photographic realism. However, if the content is pre-captured (i.e., not occurring in real time), the person rendered in the lenses of the VR headset cannot respond to the participant's actions or language in a real-world manner (attempts at AI or machine learning responsiveness notwithstanding). It is simply a recording of a previous moment in time. For co-presence and social interaction to be believable, however, and therefore effective in the case of volumetric capture, the footage must be streamed in real time (i.e., telepresence).
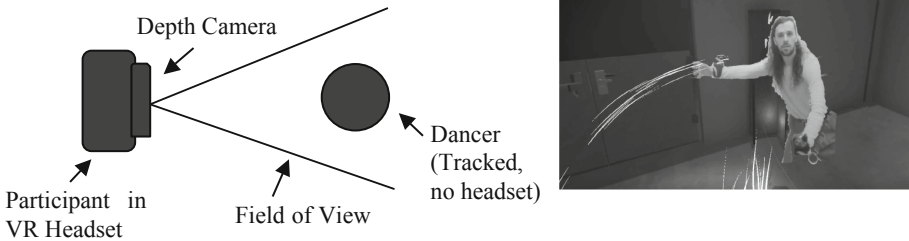
Fig. 1. Co-located Co-Presence in *Wake*: The schematic on the LEFT represents the hardware system and co-presence structure used in *Wake* for rendering the co-located dancer. On the RIGHT: The dancer (Taylor Knight) as seen in the headset, rendered in real time through the Intel RealSense camera. The dancer is tracked using Vive trackers, and can interact with virtual objects, such as the rope depicted in the image.

In order to test the possibility of real time volumetric capture, the *Wake* design probe developed a scenario in which an in-headset participant encounters and interacts with a co-present dancer who is not wearing a headset (Fig. 1). This scenario constituted a case of Asymmetric, Co-Located, Co-Present Mixed Reality (ACLCPMR). During the experience, the participant is immersed in a virtual environment rendered in the headset, interacts with both virtual and physical (tangible) objects, and engages with the dancer through simple improvised movement. In this ACLCPMR experience, the dancer is tracked in the virtual environment and is also physically co-located in the same space as the participant. The dancer was rendered in real time through a custom algorithm which utilized the feed from a depth camera (Intel RealSense) mounted on the front of the participant's headset, at the position of their eyes. Therefore, when the participant looked at something, the camera saw what they saw. Using a depth filtering algorithm, the camera feed was manipulated to only render objects which were at a certain distance from the camera (to render the dancer but not the walls of the room around them).

In *Wake*, co-presence was ultimately achieved not through simulation, as previous studies have utilized [2], but by the real time rendering of a co-located, co-present dancer. This method therefore successfully circumvented the uncanny valley problem, not by directly addressing the problems of photographic and behavioral realism in an avatar, but by side-stepping them altogether. Qualitative findings from the semi-structured phenomenological interviews revealed that a vast majority of the participants (24 out of 25) felt that they were able to make and hold eye contact with the dancer as though it was real and mutual, despite knowing that the dancer was not able to see their eyes (because of the headset). The findings in *Wake* therefore plausibly suggest that

real-world, genuine social interactions are possible within hybrid media environments like VR. More importantly, they suggest that ACLCPMR not only stands as a viable method for achieving experiential optimization when it comes to representing others in VR, but also for testing, researching, and developing new forms of human-machine mediated communication.

6 Conclusion

The standard approach to VR research and development tends to value physicalist achievements (e.g., physically realistic simulations) while overlooking both the pragmatic and the phenomenological dimensions of perceptual experience. Because the physicalism inherent in the standard approach has been inherited, and is therefore tacitly presupposed, it has the tendency of propagating itself in the literature without being subjected to critical scrutiny. As a consequence, the pragmatic and phenomenological dimensions of perceptual experience in VR can only ever remain overlooked and undervalued. The main ambition of this article has been but to disrupt this implicit propagation of presuppositions, and our decision to engage with cognitive science has been motivated precisely by this reason. More specifically, in having turned to enactivism, we have articulated an alternative, action-predicated approach to VR, one that is (1) non-reductive with respect to subjective experience, and which (2) honors the pragmatic dimension of perceptual reality. Having also illustrated the various methodological implications of our action-predicated approach, namely as regards with POEs, VOEs, immersion, flow, on- and offboarding, and experiential optimization, we believe that the next step in this line of work is to develop more rigorous empirical methods by which to test the theoretical claims and qualitative observations made in this article. Suffice it to say, though, if with this article we have at least managed to raise some interesting questions, then that alone constitutes a worthwhile beginning.

References

1. Ahmed, S.: The Cultural Politics of Emotion, 2nd edn. Routledge, London (2014)
2. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.: Interpersonal distance in virtual environments. Pers. Soc. Psychol. Bul. **29**, 1–15 (2003)
3. Biocca, F.: The cyborg's dilemma: progressive embodiment in virtual environments. J. Comput.-Mediat. Commun. **3**(2), 1–31 (1997)
4. Bulu, S.T.: Place presence, social presence, co-presence, and satisfaction in virtual worlds. Comput. Educ. **58**(1), 154–161 (2012)
5. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. HarperCollins e-books (1990)
6. Cherniak, C.: Minimal Rationality. MIT Press, Cambridge (1986)
7. Jerome, C.J., Witmer, B.: Human performance in virtual environments: effects of presence, immersive tendency, and simulator sickness. In: Proceedings of the Human Factors and Ergonomics Society: 48th Annual Meeting, New Orleans, Louisiana, pp. 2613–2617. Human Factors and Ergonomics Society (2004)

8. Mütterlein, J.: The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In: Proceedings of the 51st Hawaii International Conference on System Sciences, Waikoloa Village, Hawaii, pp. 1407–1415. Hawaii International Conference on System Sciences (2018)

9. Schwartz, R., Steptoe, W.: The immersive VR self: performance, embodiment and presence in immersive virtual reality environments. In: Papacharissi, Z. (ed.) A Networked Self and Human Augmentics, Artificial Intelligence, Sentience, pp. 108–116. Routledge, New York (2018)

10. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence **7**(3), 225–240 (1998)

11. Peterson, J.: Maps of Meaning. The Architecture of Belief. Routledge, London (1999)

12. Varela, F., Thompson, E., Rosch, E.: The Embodied Mind: Cognitive Science and Human Experience. The MIT Press, Cambridge (1991)

13. Thompson, E.: Mind in Life: Biology, Phenomenology, and the Sciences of Mind. The Belknap Press of Harvard University Press, Cambridge (2007)

14. Hovhannisyan, G., Dewey, C.: Natural & normative dynamical coupling. Cogn. Syst. Res. **43**, 128–139 (2017)

15. Hovhannisyan, G.: Humanistic cognitive science. Humanist. Psychol. **46**(1), 30–52 (2018)

16. Metzinger, T.: Phenomenal transparency and cognitive self-reference. http://www.philosophie.uni-mainz.de/metzinger/publikationen/Phenomenal_transparency_PCS2003.pdf . Accessed 24 Jan 2019

17. Vervakae, J., Ferraro, L., Herrera-Bennett, A.: Flow as Spontaneous Thought: Insight and Implicit Learning. Oxford Handbooks Online (2018)

18. Henson, A.: We're In This Together: Embodied Interaction, Affect, and Design Methods in Asymmetric, Co-Located, Co-Present Mixed Reality. Masters Thesis, Carnegie Mellon University, Pittsburgh (Pa.) (2019)

19. Chalmers, D.: Facing up to the hard problem of consciousness. http://consc.net/papers/facing.html Accessed 29 Jan 2019

20. Dreyfus, H.: Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action. Oxford University Press, Oxford (2014)

21. Dourish, P.: Where the Action Is: The Foundations of Embodied Interaction. The MIT Press, Cambridge (2001)

22. Plutchik, R.: A General Psychoevolutionary Theory of Emotion. Academic Press, New York (1980)

23. Hogarth, R.: Educating Intuition. University of Chicago Press, Chicago (2001)

24. Topolinsky, S., Reber, R.: Gaining insight into the "aha" experience. Curr. Dir. Psychol. Sci. **19**(6), 402–405 (2010)

25. Zadra, J.R., Clore, G.L.: Emotion and perception: the role of affective information. Wiley Interdiscip. Rev. Cogn. Sci. **2**(6), 676–685 (2011)

26. Haraway, D.: Situated knowledges: the science question in feminism and the privilege of partial perspective. Fem. Stud. **14**(3), 575–599 (1988)

27. Juarrero, A.: Dynamics in Action: Intentional Behavior as a Complex System. MIT Press, Cambridge (1999)

28. Hulswit, M.: From Cause to Causation: A Peircian Perspective. Kluwer Academic Publishers, Dordrecht (2002)

29. Peterson, C., Park, N., Seligman, M.E.P.: Orientations to happiness and life satisfaction: The full life versus the empty life. J. Happiness Stud. **6**, 25–51 (2005)

30. Peterson, J.: The meaning of meaning. In: Wong, P., et al. (eds.) The Positive Psychology of Meaning and Spirituality. INPM Press, Vancouver (2008)

31. Reichardt, J.: Robots: Fact, Fiction, and Prediction. Thames and Hudson, London (1978)

32. Giorgi, A.: Phenomenology and the foundations of psychology. In: Arnold, W. (ed.) Nebraska Symposium on Motivation, pp. 281–348. University of Nebraska Press, Lincoln (1976)
33. Weber, A., Varela, F.: Life after kant: natural purposes and the autopoietic foundations of biological individuality. Phen. Cog. Sci. **1**, 97–125 (2002)
34. Robbins, B.: Enactive cognition and the neurophenomenology of emotion. In: Gordon, S. (ed.) Neurophenomenology and Its Application to Psychology, pp. 1–24. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7239-1_1
35. Noë, A.: Out of Our Heads: Why You Are Not Your Brain, and Other Lessons From the Biology of Consciousness. Hill and Wang, New York (2009)
36. Gallagher, S.: Enactivist Interventions: Rethinking the Mind. Oxford University Press, Oxford (2017)
37. Vervaeke, J., Ferraro, L.: Relevance realization and the neurodynamics and neuroconnectivity of general intelligence. In: Harvey, I., et al. (eds.) Smart Data: Privacy Meets Evolutionary Robotics, pp. 57–68. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-6409-9_6
38. Vervaeke, J., Lillicrap, T., Richards, B.: Relevance realization and the emerging framework in cognitive science. J. Logic Comput. **22**, 79–99 (2012)