Phenomenology and the Cognitive Sciences (2022) 21:341–375
https://doi.org/10.1007/s11097-021-09768-5

Enactivist Big Five Theory

Garri Hovhannisyan[1] · **John Vervaeke**[2]

Accepted: 23 August 2021 / Published online: 28 August 2021

Abstract

The distinguishing feature of enactivist cognitive science is arguably its commitment to non-reductionism and its philosophical allegiance to first-person approaches, like phenomenology. The guiding theme of this article is that a theoretically mature enactivism is bound to be humanistic in its articulation, and only by becoming more humanistic can enactivism more fully embody the non-reductionist spirit that lay at its foundation. Our explanatory task is thus to bring forth such an articulation by advancing an enactivist theory of human personality. To this end, we synthesize core concepts from cognitive science, personality theory, and phenomenological philosophy in order to develop an Enactivist Big-5 Theory (EB5T) of personality. According to EB5T, personality traits are dispositional tendencies for how we come to optimally grip our distinctly human worlds. Individual differences in personality are therefore reflective of stylistic differences in optimal gripping tendencies between human beings. EB5T affords a non-reductionist understanding of the immanent teleology of the autopoietically embodied human mind as a kind of full-scale optimal gripping process that is achieved along five major dimensions of personality. To the degree that these dimensions are universal, therefore, we argue that our theory offers a viable path forward in advancing enactivist cognitive science beyond the life of a cell and into the mind of a person, a longstanding hope and ambition held by proponents of the enactive approach.

Keywords Enactivism · Five-Factor Model · Big Five · Relevance Realization · Optimal Grip · Personality Function

Garri Hovhannisyan dedicates this article in loving memory of his grandmother, Dr. Nellie Baghdasarian. This article draws on research supported by the Social Sciences and Humanities Research Council of Canada.

✉ Garri Hovhannisyan
garri.hovha@gmail.com

[1] Department of Psychology, Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282, USA

[2] Department of Psychology, University of Toronto, Toronto, Canada

1 Introduction

Profound philosophical insights have the potential to not only elucidate the limits of what we already know but to radically alter our very horizon of understanding in the process. In the early 1990s, Varela et al. (1991) sought to insightfully transform the theoretical and conceptual landscape of cognitive science through their presentation of the "enactive approach" to cognition. The guiding idea of enactivism was that cognition is not a skull-bound process of symbol manipulation in relation to a pregiven world but rather a procedural and participatory act through which all meaningful worlds are brought forth (Noë, 2009; Thompson, 2007; Varela et al., 1991).

In the three decades since its initial appearance, enactivism has burgeoned into a comprehensive research program with a well-defined set of philosophical and scientific aims. The distinguishing feature of the enactive approach is arguably its commitment to non-reductionism and its philosophical allegiance to first-person approaches, like phenomenology, despite its simultaneous rootedness in the natural and life sciences (Fuchs, 2018; Gallagher, 2017; Thompson, 2007; Varela, 1996; Varela et al., 1991). For example, enactivists have generally sought to secure scientifically plausible understandings of such phenomena as consciousness, meaning, and organismic teleology without resorting to strictly mechanistic explanations of the type characteristic of earlier cognitive science frameworks (Bockelman et al., 2013; Di Paolo, 2005; Hovhannisyan, 2021; Olivares et al., 2015; O'Regan and Noë, 2001; Robbins & Gordon, 2015). To this end, the non-reductionist attitude of the enactive approach has been exemplified and effectively implemented in neurophenomenological research contexts with various subject matters like time-consciousness (Varela, 1999), emotion (Robbins, 2013), perception (Lutz & Thompson, 2003), and even the experience of seizures (Petitmengin et al., 2007), thereby attesting to enactivism's overall promise and potential as a genuinely scientific enterprise.

Our article begins with the idea that the enactive approach is bound to become more distinctly human in its articulation as it continues to develop in line with its non-reductionist aims (Fuchs, 2018; Gordon 2013; Hovhannisyan, 2018, 2021; Robbins, 2013; Robbins & Gordon, 2015; Sundararajan & Ho, 2020; Thompson, 2004, 2005, 2007, 2011; Varela et al., 1991). Our explanatory task is to bring forth such an articulation by advancing an enactivist theory of human personality. The reasons for this undertaking are twofold. The first reason is that enactivist accounts seek to be comprehensive and non-reductionist, yet they typically employ minimal models of cognition (e.g., autopoietic sense-making or sensorimotor perception) as their basic reference points when attempting to understand cognitive life at the level of human realization. Any inferences made about human minds on the basis of single-celled organisms or sensorimotor activity, however, are phenomenologically limited and should be drawn with a degree of theoretical caution. Bridging the phenomenological gap between minimal and human cognition requires the articulation of more sophisticated models of cognition, not merely as instantiated in the domains of cellular (autopoietic) or sensorimotor

life, but as lived in the uniquely human dimension of experience. We believe that an enactivist theory of personality is necessary for bridging this gap by offering the kind of theoretical model that is needed to move (enactive) cognitive science beyond the life of a cell and into the mind of a person.

The second reason is that there is also an a priori need for extending enactivism into the domain of personality research—insofar as questions of personality are central to the functionality of the human mind and enactivism is concerned with articulating a *comprehensive* understanding of human cognitive life. Recent work by DeYoung (2015) has already initiated a dialogue between cognitive science and personality theory. The theoretical importance of DeYoung's interdisciplinary contribution cannot be overstated, since it has elevated and (arguably) for the first time made central the role of cognitive science in conceptualizing personality structure and function. Yet explicit engagement with this literature from a specifically enactivist point of view has been lacking. In articulating an enactivist theory of personality, we thus seek to also ameliorate this (second) gap in the enactivist literature by demonstrating how questions of human personality structure and function can be broached from an enactive perspective.

The approach of our article is synthetic and draws inspiration from different theoretical traditions, such as personality theory, cognitive science, and phenomenological philosophy. Our primary thesis is that personality traits are dispositional tendencies (i.e., dynamical constraints on possibilities for action existing over longer timescales) for how we come to optimally grip our distinctly human worlds. Individual differences in personality, on our account, are therefore reflective of stylistic differences in optimal gripping tendencies between human beings. In advancing this thesis, our broader philosophical intent is to endorse an alternative metaphor for the mind to that of a computer or that of an autopoietic cell (taken from cognitivism and enactivism, respectively). Specifically, we propose and claim that the metaphor of the *mind as grip* more aptly conveys the embodied, enactive, and situated aspects of consciousness and cognition with which enactive cognitive science is primarily concerned and upon which our own enactivist theory of personality is largely predicated.

We advance a discussion of the core ideas undergirding our thesis through a series of dialectical steps. Our conversation begins with the recognition that the prototypical approach to modeling cognition in enactivist terms, namely through comparing (explicitly or implicitly) the human mind to the life of an autopoietic cell or even grounding it in an account of embodied sensorimotor agency, needs to be systematically articulated into a model of human cognitive life that is informed by and in dialogue with contemporary psychological research. Such an enriched "humanistic" model is necessary, we believe, because it can more plausibly depict the domain of human experience as it is lived, and, in the process, more meaningfully align with and leverage enactivism's broader philosophical vision of a non-reductionist science of the mind.

Thus, we turn to DeYoung's (2015) Cybernetic Big-5 Theory (CB5T) of personality which, we think, is able to provide such a model by articulating an account of human cognitive life along five major dimensions of personality functioning. However, we note that a theoretical integration of CB5T and enactivism requires a

degree of mutual compatibility that has not yet been established, and we identify two requirements for making such an integration possible. First, as a theory that endeavors to explain personality function in cognitive terms, we find it important to ensure that CB5T does not succumb to the frame problem—i.e., the problem of how cognition realizes what is relevant. In our article, we discuss and treat the frame problem as a basic theoretical constraint that applies to all explanatory theories of cognitive function, including enactivist theories of cognition as sense-making or as sensorimotor activity (we explain our rationale for such a treatment in greater detail in the body of this paper). And second, due to its grounding in cybernetics, we determine the need to empty CB5T of any representational and computational meanings it may (implicitly or explicitly) harbor prior to integrating it with enactivism, meanings to which enactivism is necessarily opposed at the level of basic presuppositions.

In order to meet both of these requirements, we first position ourselves within a recent literature that promises to circumvent the frame problem in a non-representational and non-computational manner—a necessary condition for affording theoretical compatibility with the enactivist paradigm. The theoretical framework that is articulated in this literature explains that cognition realizes relevance (i.e., circumvents the frame problem) through processes of optimization involving trade-off relationships between opponent yet complementary learning strategies with respect to a mutual goal (Hovhannisyan & Dewey, 2017; Vervaeke & Ferraro, 2013; Vervaeke et al., 2012). For example, cognition can realize relevance with respect to the goal of "evading threats" by dynamically optimizing the trade-off between "fight" and "flight" in an opponent-processing fashion. This framework thus explains cognition in terms of relevance realization (RR) and, accordingly, identifies RR as the mark of the cognitive.

Upon explicating the cognitive architecture articulated by the theory of RR and reviewing its explanatory edge over representational and computational accounts of mind, we argue that RR forms the dynamical basis of Merleau-Ponty's notion of "optimal grip." And, upon construing it in enactive, embodied, and situated terms qua a reformulation of RR *as* optimal grip, we suggest that RR has the potential to explain important aspects of the immanent teleology of the autopoietically embodied *human* mind, which is of central concern to enactivist cognitive science.

In the final stage of our argument, we proceed to demonstrate how CB5T can be integrated with enactivism to produce an Enactivist Big-5 Theory (EB5T) of personality. We thus ground CB5T in RR in order to afford an understanding of traits as dispositional tendencies for optimally gripping our distinctly human worlds. This grounding relation allows us to conceive of human personality functioning in enactive terms but in a way that importantly extends enactivist understandings of human cognitive life beyond the minimal models employed previously. This is because optimal gripping (at least on our account) is an embodied process that is achieved through the dynamics of opponent-processing and—much like in the classical autopoietic model—it is a process that (at least in principle) does not entail any representational or computational commitments. However, unlike the autopoietic model, EB5T articulates an enriched, multidimensional, humanistic account of cognitive activity that has the potential to elucidate human experience on its own terms by way of its integration with important findings from contemporary personality

research. In addition, we find that our enactive approach to personality affords a yet unexplored discussion of extended cognition whereby EB5T can explain not only optimal grip within individual cognition but also optimal grip between people in distributed cognition. Specifically, EB5T thematizes the adaptive function of personality type variation within populations of persons, hence extending the continuity of life and mind to the domain of personality and even culture.

In the conclusion, we discuss the implications of our enactivist theory of personality for cognitive science research, personality theory, and clinical psychopathology and psychological assessment practice. We conclude by summarizing that our presentation of EB5T is not meant to be a theoretical replacement for CB5T, but an alternative that may further facilitate the interdisciplinary and much-needed dialogue between cognitive science, personality research, and clinical psychological practice—that is, by laying a path for those who are wishing to engage with questions of human personality from a specifically enactivist point of view.

2 From minimal to human cognition

Radical ideas are never born in a vacuum but emerge in response to fundamental philosophical problems being confronted at the time of their inception. It is thus important not to forget that the enactive approach was originally presented as a "continuation of a program of research founded over a generation ago by the French philosopher, Merleau-Ponty" (Varela et al. 1991, p. xv). Yet over a decade after the initial publication of *The Embodied Mind*, one of the authors, Evan Thompson (2004), made an important observation: "although the ideas about embodied cognition in this book have been widely acknowledged and assimilated by the field, the book's central theme has yet to be fully absorbed" (p. 382). He then poignantly elaborated that this "theme is the need for back-and-forth circulation between scientific research on the mind and disciplined phenomenologies of lived experience" (2004, p. 382).

Unfortunately, we find that Thompson's observation still rings true in the present-day context, as some of the truly radical ideas that played a formative role in organizing early enactivist thinking have since been "relegated to the fringes of discussion, if not completely ignored" (Vörös, 2020, p. 92). Our purpose in this section is to demonstrate why we think this to be the case and how we believe enactivism ought to proceed in order to revivify and more fully embody the radical philosophical spirit that originally animated it into being. Simply stated, our critical assessment suggests that enactivism's main theoretical and philosophical strategy for securing non-reductionism with respect to human lived experience risks falling into a kind of (paradoxical) self-contradiction, and that this self-contradiction can only be ameliorated by developing the enactivist understanding of the embodied mind beyond its present state through a disciplined dialogue (i.e., a back-and-forth circulation) with the empirical human sciences. Let us begin with a discussion of the nature of this paradox and its context before we elaborate on what makes such a dialogue necessary to begin with.

Whereas most traditional (cognitive scientific) accounts of human lived experience have ventured to explain how phenomenal states in the mind are caused by physical states in the brain, enactivism begins by adopting a "radical embodiment" formulation of the mind–body problem as a "body-body problem" (Thompson, 2004, p. 385). Thompson (2011) explains that such a reformulation finds its roots in the phenomenological tradition first articulated by Husserl (1970) and later developed by Merleau-Ponty (1942/1963), which "distinguishes between two ways the body can be disclosed to our experience—as a material thing (*Körper*) and as a living subject of experience or lived body (*Leib*)" (p. 8, emphasis in original). A phenomenological treatment of embodiment demonstrates that—contra classical (reductionist and dualist) accounts—the objective, living body (*Körper*) and the subjectively lived body (*Leib*) are not fundamentally opposed ontological domains or properties. Rather, they are just two distinct thematizations of one and the same thing (the body) in our field of experience—as Thompson (2011) construes it, "two types within one typology of embodiment" (p. 8). As such, the task of the body-body problem becomes to "understand the relation between the body as a living being and the body as a lived body or bodily subject of experience" (p. 8). Consequently, if the mind is embodied and embodiment can be thematized in two distinct yet interrelated ways, then in virtue of the body-body formulation, "the [explanatory] gap is no longer absolute because in order to formulate it we need to make common reference to life or living being" (Thompson, 2011, p. 9).

The enactive approach then advances a second proposition, the "deep life-mind continuity thesis," to further leverage its non-reductionist position (Kirchhoff & Froese, 2017; Thompson, 2007). According to the deep continuity thesis, the core organizational and phenomenological properties of complex minds (e.g., human intentionality) are an articulation of those found in basic forms of life (i.e., organismic teleology) (Thompson, 2007). In order to begin bridging the conceptual gap between the living, biological body and the lived, phenomenological body, enactivism subsequently deploys the theory of autopoiesis to thematize the core set of (organizational and phenomenological) properties that are said to be shared by both life and mind (Thompson, 2007, p. 129; 2011). Autopoiesis seeks to secure a necessary and sufficient definition of what it means for a system to be a living system and promises to explain the adaptive (teleological) character of biological systems through means of self-organized dynamics (without, therefore, appealing to internal representations of an external world). To this end, enactivism takes the single cell as the paradigmatic example of an autopoietic system in order to begin its articulation of mindedness. Although the classical autopoietic model has been deemed insufficient for explaining mind—partly because it subsumes systems that we would not normally consider to be either living or minded (e.g., tornadoes and convection currents)—the growing consensus in the literature, with which our own view is aligned, is that autopoiesis *and* adaptivity are jointly sufficient for capturing the organizational dynamics and intrinsic teleology of embodied minds (Bourgine & Stewart 2004, Di Paolo, 2005; Hovhannisyan, 2021).

This line of argument affords the formulation of two major theoretical notions or principles commonly employed by enactivists. The first is an account of minimal mindedness or cognition that is instantiated at the level of the single-celled

(autopoietic) organism and whose form is thus necessarily (biologically) embodied, enactive (qua sensorimotor activity), embedded (i.e., pragmatically situated), and extended (i.e., world-involving)—and, most recently, affective (i.e., emotive; see Gallagher, 2017). The second, accordingly, is an articulation of the first and is employed by enactivists when theorizing about the structures, dynamics, and phenomenology of *human* lived experience. Specifically, it is the idea that *if* human intentionality is an articulation of organismic teleology, then *insofar as* the deep life-mind continuity thesis holds, the immanent teleology of the autopoietically embodied human mind is *like* that of the cell: it is also necessarily embodied, enactive, embedded, extended, and affective. Consequently, to the degree that such a relation holds, one can begin to reason about the human mind as one would a living cell.

What we have provided here is not a detailed summary of the enactive approach but a rough sketch of its basic logic that was meant only to demonstrate three important points. The first is merely descriptive and states that enactivism is, for the most part, fundamentally non-reductionist with regards to experience in virtue of its adoption of the radical embodiment reformulation of the mind–body problem (i.e., the body-body problem). The second point, also descriptive, reiterates that in spite of its non-reductionism, enactivism is nevertheless naturalistically styled insofar as it grounds mind in life by endorsing an understanding of lived experience as *biologically embodied* qua the deep life-mind continuity thesis and the theory of adaptive autopoiesis. The third point, however, is critical and states that despite securing (at least in principle) a naturalistic *and* non-reductionist approach to understanding human lived experience, unless enactivism can move beyond the autopoietic cell model by developing more sophisticated models of human cognitive life, it risks reducing or precluding the phenomenological properties of human lived experience from its understandings of human cognitive life—this is because, after all, (1) the relation between human intentionality and organismic teleology is one of *continuity* and not *identity* and (2) the degree to which one is "like" the other is not concretely specified (for a similar discussion, see Froese & Di Paolo, 2009). The irony that is revealed through this critique is that, in order to evade reductionism, enactivism has to appeal to the core organizational properties of basic forms of life (organismic teleology). Yet in needing to endorse a deep continuity of life and mind in the absence of a phenomenologically plausible model of human cognitive life, enactivism risks overfitting the single-celled autopoietic model of cognition in understanding the obviously more complex dynamics of the autopoietically embodied *human* mind—both in its living and its lived dimensions.

Fundamentally, therefore, we find that enactivism is caught in a circle of its own making. For it was brought to its current position in virtue of its original commitment to non-reductionism (i.e., to examine human lived experience on its own terms). Yet in its present state, enactivism paradoxically risks committing a different kind of reduction which, too, it surely wishes to avoid: reducing the human mind to the life of a cell. We do not think that this circle is inherently inescapable or that it dooms the enactivist project to failure as a genuinely non-reductionist science of lived experience. Rather, we see it as an indication that it is time for enactivism to evolve into a *humanistic* cognitive science whose focal aim is to articulate the structures and dynamics of human lived experience on their own terms. The only

reasonable way out of this circle, we believe, is to extend enactivism beyond its current theoretical boundaries and into the domain of the empirical human sciences, specifically personality theory, through the kind of back-and-forth circulation that Thompson deemed part and parcel of the original enactivist vision. In doing so, we feel that enactivism can at once become more distinctly human in its articulation while more fully embodying the spirit of non-reductionist science that lay at its foundation. In the following section, we review DeYoung's (2015) cybernetic account of human personality, CB5T, which we believe can afford the kind of humanistic turn that we are seeking to initiate with our article as we work toward our enactivist theory of personality.

3 Cybernetic big five theory

Our discussion must begin with the five-factor model (FFM) of personality, or "the Big Five," which forms the theoretical basis of CB5T. The Big Five is arguably one of the most well-established and widely used models of personality for predicting human behavior (e.g., McCrae & Costa, 2003). Unlike most other models of personality (e.g., psychodynamic, humanistic, social-cognitive, etc.), the origins of the Big Five lie in a lexical approach, meaning that its conception of human personality structure is empirical rather than rational in its foundation (McCrae & Costa, 2003). The robustness and universality of the Big Five is evident in that it has been cross-culturally validated and shown to be a reliable measure of personality across time (McCrae & Costa, 2003; McCrae et al., 2010). Moreover, its predictive validity makes it a highly valuable and desired tool for many different disciplines in psychology, including organizational psychology (e.g., Hurtz & Donovan, 2000; Major et al., 2006; Zhao & Seibert, 2006), clinical psychology (e.g., Costa & McCrae, 1992; John et al., 2008; Widiger & Costa, 2013), and others (e.g., Donnellan et al., 2004; Komarraju et al., 2011; O'Connor & Paunonen, 2007).

Despite its predictive utility as an assessment tool, a fundamental limitation of the Big Five is that it describes but does not *explain* human behavior (Fleeson & Jayawickreme, 2015). As a result, any attempt to account for a given behavior in terms of the traits described by the Big Five is bound to confront circularity issues, since traits are themselves defined in terms of behaviors. Given that the central aim of personality research is to predict and, ultimately, explain human behavior (Cervone & Mischel, 2002), it is important for a theory of personality, such as the Big Five, to be able to explain personality structure and function in causal terms without succumbing to such circularity issues in the process.

DeYoung's (2015) seminal article, entitled "Cybernetic Big Five Theory" (CB5T), promises to circumvent this explanatory challenge by grounding the Big Five in cybernetics, "the study of goal-directed, self-regulating systems" (p. 33). CB5T presents a comprehensive theoretical framework for understanding the structural–functional organization of human personality along five major dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. More specifically, it provides an integrative, explanatory account of the adaptive functions of traits in their cognitive and evolutionary dimensions,

moving beyond the naïve behavioral definitions of more classical accounts and evading circularities of the kind mentioned above. Since its initial publication, DeYoung has further built upon his framework, such as by developing a cybernetic account of psychopathology based on CB5T (DeYoung & Krueger, 2018; DeYoung & Weisberg, 2018) and even articulating an account of the functional neurobiology underlying its structure and function (Allen & DeYoung, 2016; DeYoung, 2010, 2013; DeYoung et al., 2010).

Because of its grounding in the Big Five, which are said to describe *universal* human structures, we believe that initiating a disciplined dialogue between CB5T and the enactive approach can provide the latter with the theoretical ingredients that it needs in order to move beyond the life of a cell and into the mind of a person. We thus regard the initiation of such a dialogue as a necessary step toward accounting for the immanent teleology of the autopoietically embodied human mind while exemplifying the very non-reductionist attitude around which enactivism is fundamentally organized.

CB5T is the consequence of a careful theoretical synthesis of key ideas drawn from cybernetics, evolutionary theory, and the Big Five theory of personality, as well as functional neurobiology and personality neuroscience (Allen & DeYoung, 2016; DeYoung, 2013, 2015). However, the basic building blocks of personality are "traits," which should be thoroughly defined before the structural and functional elements of personality are even discussed. In his article, DeYoung (2015) adopts the definition of traits as "probabilistic descriptions of relatively stable patterns of emotion, motivation, cognition, and behavior, in response to classes of stimuli that have been present in human cultures over evolutionary time" (p. 35). Additionally, since traits describe a person's situated dynamics over longer timescales, they act as attractors for the person's moment-to-moment dynamical unfolding and are therefore predictive (in a probabilistic, not deterministic sense) of what states the person is likely to inhabit at any particular point in time (DeYoung, 2015, p. 35). Such a definition has great theoretical utility for at least three reasons. First, it formulates an understanding of traits in dynamical systems theory (DST) terms, to which enactivist cognitive science is particularly sympathetic and theoretically amenable. Second, in so doing, it establishes a kind of continuity between personality and cognition that we require for our own theoretical project in this paper: personality is just cognition considered from a longer timescale. Rather than claiming a strict division between higher-level psychological phenomena (i.e., personality traits) and lower-level psychological phenomena (e.g., cognitive states, emotional states, motivational states, etc.), this definition locates the relevant difference between these two levels as a function of the timescales from which each level considers the person's situated dynamics. Thus, traits refer to the unfolding of a person's dynamics over longer timescales (e.g., spanning months or years), whereas states refer to the unfolding of a person's dynamics over shorter timescales (e.g., spanning minutes, hours, or days). In both cases, the person's activity consists of the same "stuff" (i.e., cognitive, emotional, motivation, behavioral, etc., patterns of activity), and, so, the question of how to relate two categorically different phenomena (states versus traits) does not necessarily arise. Finally, it implies that, as patterns of activity that evolved *in response* to

"classes of stimuli that have been present in human cultures over evolutionary time," traits are not only dynamical processes but are also functional or adaptive.

Thus, DeYoung (2015) turns to cybernetics, "the study of goal-directed, self-regulating systems" which offers him an account of normativity by which to explain the various functions of the Big Five. In particular, DeYoung ventures to explain the general functionality of the human personality as a cybernetic system by linking the functions underlying each of the individual traits to different stages of the cybernetic cycle: (1) goal activation, (2) action selection, (3) action, (4) outcome interpretation, and (5) goal comparison (DeYoung, 2015). In the author's words, "each of the five traits corresponds to interpersonal variation in one of the major functional categories of intrapersonal mechanism [the Big Five] involved in the operation of the human cybernetic system" (DeYoung, 2015, p. 41). CB5T, in other words, conceptualizes traits as parameters on the system's cybernetic functioning with respect to set problem domains ("broad class of stimuli," in DeYoung's terms) for which the traits have adapted, such that interpersonal differences in personality are essentially attributable to variations in the parameters operant on each trait's functioning at the intrapersonal level.

DeYoung's cybernetic approach to personality, CB5T, is deeply insightful and integrative in how it transforms the Big Five from a descriptive theory into a comprehensive *explanatory* framework for personality. However, it faces two important limitations. First, as DeYoung (2015) himself admits, the Big Five do not neatly line up with the various stages of the cybernetic cycle but seem to be differentially involved in all stages at different times (p. 34). The reason for this is that the five stages do not occur serially but are always occurring in parallel; yet various combinations of traits always seem to be implicated at different stages of the cybernetic process. The stepwise nature of the cybernetic process is therefore at best heuristically helpful in understanding personality function, but it fails to constitute a clear, literal depiction of the functional processes implicated therein, since these processes do not in fact unfold serially but dynamically and in parallel. Thus, the exact dynamical nature of these processes is not specified, and their precise conditions of unfolding remain unclear. To the degree that any functional account of personality is grounded in cybernetics in such a manner, therefore, its value as a functional theory is going to be limited, even if it is heuristically useful.

The second limitation is that cybernetics is deeply embedded in a theoretical paradigm that predates enactivist cognitive science (Froese, 2011), such that, when taken literally, harbors some classically representational and computational commitments in its account of how the cybernetic cycle is achieved (i.e., as a stepwise process that recycles in an algorithmic fashion). This is undoubtedly problematic for our theoretical aims because enactivism by definition tends to be anti-representational and anti-computational in its commitments. Thus, if CB5T is to be integrated with enactivism, it must be ensured that CB5T does not explicitly or implicitly presuppose any such commitments in its formulation of personality functioning. However, we must note that as a *cognitive* theory of personality function, it must also be ensured that CB5T does not succumb to the notorious frame problem. The frame problem is the problem of explaining how cognition realizes what is relevant and is arguably a basic theoretical constraint that applies to all explanatory accounts

of cognitive function (Dennett, 1987; Vervaeke et al., 2012). The immanent problem here for CB5T, therefore, is that representational and computational accounts of cognitive function have been shown to *necessarily* succumb to the frame problem in virtue of their very commitments (Dennett, 1987; Vervaeke et al., 2012), meaning that slipping representational or computational meanings into a cognitive scientific account of personality is to be avoided not only to ensure compatibility with enactivism but also, indeed, to evade the frame problem. In the following section, we discuss the importance of the frame problem for cognitive science, in general, and the (enactive) cognitive science of personality, in particular, as we review an account of cognition that promises to circumvent the frame problem in a manner that fits with the theoretical aims of our article.

4 Relevance realization

In order to circumvent the theoretical difficulties facing CB5T that were outlined in the previous section while ensuring theoretical compatibility between enactivism and CB5T, we devote this section to reviewing a recent theory of cognition as relevance realization (RR) that promises to circumvent the frame problem through non-representational, non-computational means. The guiding theme of this section is that RR forms the dynamical basis of Merleau-Ponty's notion of optimal grip and therefore constitutes an inherently enactive, situated, and embodied account of cognition that can be used to transform CB5T into an Enactivist Big-5 Theory (EB5T) of personality.

Any theory of cognitive function must ultimately explain how it is possible for cognitive agents to effectively zero-in on relevant information, to put a "frame" around what they consider to be important for functioning (Dennett, 1987; Vervaeke et al., 2012). The so-called "frame problem" is a fundamental problem for cognitive science and must be circumvented if cognition is to be explained. The pervasive difficulty is that cognitive agents are bombarded with a virtually indefinitely large set of information despite the fact that their means of coping with information are severely limited both in terms of time and cognitive resources (Cherniak, 1986; Vervaeke et al., 2012). Consequently, the only way to guarantee that one has realized what is relevant is by conducting an exhaustive search of the problem space, but this is impossible since any attempt to cope with an indefinitely large data set with finite means will lead to combinatorial explosion and therefore entail cognitive suicide (Cherniak, 1986). The explanatory task, then, is to show how cognition is able to realize relevance while evading combinatorial explosion in its strategies for coping with the world.

Vervaeke et al. (2012) articulate two requirements that a theory of RR must meet if it is to circumvent the frame problem. First, such a theory must not commit itself to a substantive definition of relevance in explaining how cognition does RR. This is because the term "relevance" does not denote a stable or homogeneous class of entities, but a dynamical and evolving category akin to that of "biological fitness." It is not possible, therefore, to explain RR by generating a scientific theory of relevance since scientific theories make inductive generalizations and inductive

generalizations require the phenomena of study to form a homogeneous and stable class. Conversely, any attempt to account for relevance as a substantive category by creating a general-purpose learning algorithm—a typical strategy used by computational accounts of cognition—will necessarily be combinatorially explosive and thus confront the frame problem since potentially *anything* can (at least in principle) count as relevant given the right context. Hence, what a theory of cognitive function must explain is not *what* is relevant, but the stable set of processes by which cognition is able to do RR in an ongoing yet contextually sensitive fashion (Vervaeke et al., 2012).

The second requirement is that such a theory must evade homuncular fallacies, which is to say, the capacity for RR must not be tacitly presupposed in the theory's explanation of how cognition realizes relevance (Vervaeke et al., 2012). For example, cognitivist accounts typically advance some version of the claim that cognition achieves real-world problem-solving by first representing goals and pathways to solutions and second by acting on these representations to actually solve problems. *Good* representations, they maintain, are what allow cognitive agents to select possible pathways for attaining their goals, thereby solving their problems and affording functional action. What such theories fail to note, however, is that representations are aspectual in that, when one is representing a given thing, one is representing only *some* aspects of it at the exclusion of others (e.g., when I imagine an apple, I imagine it only from one side; Vervaeke et al., 2012). This implies that generating good representations requires selecting only those aspects—from the (inexhaustible) set of all possible aspects characteristic of a situation—that would be relevant for acting functionally with respect to the situation. However, the ability to determine and differentiate aspects of a situation based on their relevance presupposes the capacity to realize what is relevant as it entails the selection of an infinitesimal subset of possibilities from an indefinitely large set all while evading combinatorial explosion (Vervaeke et al., 2012). Without the capacity for RR, no representation that is generated can be guaranteed to be "good" without first employing a combinatorially explosive search strategy. All of this is to say that cognitivist accounts of the representational variety fail to account for cognitive function except by a slight of hand, since they presuppose the very thing they have set out to explain (RR). In virtue of their circularity, therefore, such accounts do not meet the second requirement and therefore fail to secure a theoretically plausible means of circumventing the frame problem.

Most theoretical accounts of cognitive function fail to meet at least one of these requirements and are thus doomed to confront the frame problem (Vervaeke et al., 2012).[1] In their seminal article, Vervaeke et al. (2012) begin articulating a theory of RR that promises to circumvent the frame problem by meeting both of these requirements, which is to say, their theory neither commits to a substantive definition of relevance nor presupposes the capacity for RR in its explanation

[1] The homuncular fallacy is not unique to representational accounts of cognitive function but applies to other important domains as well (e.g., theories of language-use, categorization, agent-environment interaction, etc.). For a systematic review of the circularity issues plaguing theories of cognitive function, see Vervaeke et al. (2012).

of cognitive function. The authors articulate the rationale behind their theory by drawing an analogy to Darwin's theory of evolution by way of natural selection. Just as Darwin's theory dispenses with the project of theorizing about biological fitness and instead explains how biological fitness is continually realized in an evolving and self-organized fashion through the mechanisms of natural selection, so too does Vervaeke et al.'s (2012) theory of RR dispense with theorizing about relevance as such; instead, it attempts to explain the mechanisms by which relevance—i.e., *cognitive fitness*—is continually realized in a self-organized, evolving, and dynamically situated fashion. We may refer to Vervaeke et al.'s (2012) approach to the frame problem as an "opponent-processing" approach that replaces the computational metaphor for mental functioning with the metaphor of the mind as a dynamically situated cognitive economy. The general principles behind their theory can be illustrated in the following six steps:

1. First, specify some problem domain for the agent that constitutes a condition for successful interaction with the world (e.g., "avoiding harm").
2. Second, specify two distinct (heuristic) strategies that can be used in the service of accomplishing the same goal but in complementary ways (e.g., "fight when faced with a threat, but risk being eliminated now" versus "flee when faced with a threat, but risk being eliminated later").
3. Third, pair these opponent strategies in a trade-off relationship and set them to optimize for reward relative to the given problem domain (e.g., avoid danger) in a dynamically recursive fashion (e.g., set the "fight-or-flight response to potentially dangerous stimuli" to optimize for threat-avoidance through an iterative process across a variety of contexts).
4. Fourth, repeat steps 1–3 for other sets of problems until you have articulated a *problem landscape* involving a complex cognitive economy of competing goals and problems (e.g., hunger, thirst, sleep, etc.)—and concomitant opponent-processes—facing the agent.
5. Fifth, specify processes for prioritizing the importance of different, competing goals—mechanisms which can also be realized by means of opponent-processing so as to make optimization with regards to prioritization possible (e.g., how to prioritize the relevance of competing motivational goals like thirst, hunger, sleep, etc. in a self-organized fashion).
6. Finally, set the agent's cognitive economy to optimize for reward *across* various problem domains *and* various timescales, so that the scale-invariant optimization of the agent's total economy eventually and flexibly results in the realization of what is "relevant" in a context-sensitive, evolving, and self-organized fashion in relation to the agent's situated interaction with the environment (i.e., the problem landscape).

This stepwise formulation of the RR framework ensures that a substantive definition of relevance is not assumed and that RR is not tacitly presupposed in its explanation of cognitive function—and, so, circumvents the frame problem. The first requirement is met by framing relevance not as a fixed value specified to the

agent from the top-down, but as an emergent value based on the ongoing optimization of the trade-off between complementary *heuristic strategies* relative to a problem landscape across different timescales (i.e., a complex cognitive economy of competing goals, problems, and opponent-processes). Accordingly, by placing complementary *heuristic strategies* (which, contra *algorithms*, conduct a partial rather than exhaustive search of the problem space) in *opponent* pairs, the process of RR evades combinatorial explosion whilst nevertheless retaining a range of cognitive flexibility that allows it to adapt to its conditions by alternating between its complements in a self-organized manner.[2]

In their article, Vervaeke et al. (2012) establish that there are three opponent pairs that act as constraints on the functioning of the cognitive economy as a whole and enable it to continually do RR relative to its problem landscape in a non-homuncular fashion. The first constraint is Cognitive Scope (CS), which optimizes for the applicability of the system's information to its world by trading-off the demand for special-purpose learning with that of general-purpose. Special-purpose learning strategies are metabolically costly since they "process information in a particular domain with high rigor (with the possibility of finding reward)," but they "risk loss to the cognitive system since […] their domain may not yield sufficient reward to break even" (Hovhannisyan & Dewey, 2017, p. 133). On the other hand, general-purpose strategies are "metabolically cheap […] since they process information across several domains with low rigor," but they "risk high opportunity cost […] by forfeiting rigorous examination of domains that do yield high rewards" (Hovhannisyan & Dewey, 2017, p. 133). Consequently, "hard commitment to either of these strategies (general purpose or specific) is undesirable" (Vervaeke et al., 2012, p. 91). However, by pairing them in a trade-off relationship, a system can optimize for applicability such that the risks of each strategy can be continually mitigated by the other in a self-organized manner.

The same logic applies to the second constraint, Cognitive Tempering (CT), which optimizes for the projectability of the system's information by pairing exploitative and explorative learning strategies together (Hovhannisyan & Dewey, 2017; Vervaeke et al., 2012). On the one hand, exploitative learning strategies aim to maximize reward by pursuing opportunities that can be "easily and immediately accessed" (Hovhannisyan & Dewey, 2017). The obvious risk, of course, is that such strategies can forego meaningful opportunities that are to be found elsewhere. Explorative learning strategies, on the other hand, aim to mitigate this risk by pursuing more (temporally and spatially) distal opportunities instead; but, in so doing, they "risk losing out on easily and immediately accessible opportunities at a high cost (the risk that exploitative strategies mitigate)" (Hovhannisyan & Dewey, 2017). Similar to CS, therefore, CT realizes relevance as the risks of one strategy continue

[2] Asserting the presence of various problem domains or heuristics does not necessarily render this theoretical formulation homuncular, since the theory of RR is not concerned with explaining the causal origin of the particular problems confronting agents (which are likely to be evolutionary and biological in their origin), but rather how cognitive agents can adaptively solve the problems which they are already dealt with while evading combinatorial explosion.

to be mitigated by the implementation of the other strategy in the pair in a complementary and self-organized manner.

Unlike CS and CT, however, the final constraint, Cognitive Prioritization (CP), concerns "the structure and prioritization of [various] cost functions," rather than "how [individual] cost functions might be heuristically optimized" (Vervaeke et al., 2012, p. 93). A system is "flexibly gambling" its resources when optimizing for CP, namely, by trading-off the complementary strategies of focusing and diversifying:

> CP is achieved through the interaction of several competing cost functions, each of which is an internal metric that tracks progress in its respective problem domain (e.g., finding food in the environment, avoiding predators in the environment, etc.). When low positive feedback is received in one domain (i.e., low energy reserves), the system's resources become more focused on solving that particular problem (i.e., finding food). Once the issue has been resolved, though, the system's resources become less focused on that one problem, and become more diversified instead. (Hovhannisyan & Dewey, 2017, p. 133)

At this point, steps one through six of Vervaeke et al.'s (2012) theory of RR have all been elaborated on and explained. According to RR, the mind is best construed not as a computational machine whose behavior is mediated through internal representations, but a dynamically situated cognitive economy that realizes relevance by optimizing for reward across various problem domains (applicability) and timescales (projectability) by means of interrelated opponent-processes (CS, CT, CP).

Importantly, the authors further theorize about two emergent, higher-order (superordinate) constraints that govern the overall cognitive economy from the top-down and unify the various lower-order (subordinate) constraints into a single, complex dynamical system: Efficiency and Resiliency. The higher-order constraint of Efficiency fulfills a selective function on the economy by reducing *entropy* or variation in the possibilities for action available to the cognitive agent. Efficiency thus subsumes general-purpose learning, exploitative learning, and focusing. Although less metabolically costly, too much Efficiency risks the loss of long-term adaptability to eventual and ongoing situational change. In this regard, the higher-order constraint of Resiliency fulfills a complementary, enabling function by increasing *entropy* or variation in the possibilities for action available to the cognitive agent by subsuming special-purpose learning, explorative learning, and diversification. Resiliency is necessary to the degree that a cognitive system must adaptively self-organize its resources (cognitive economic constraints) to appropriate its fit to an environment that is in flux, if it is to survive. Of course, too much unwarranted variation within the economy will only lead to undue chaos and deplete the system's resources, thus causing the system too much instability to meaningfully adapt to its problem landscape. By optimizing the trade-off between Efficiency and Resiliency, therefore, the cognitive economy can continually redesign itself and re-appropriate its fit to the environment as needed, thereby realizing what is relevant in a contextually sensitive and non-homuncular fashion.

By this point, it should be clear that RR is a post-cognitivist theory of cognition. But to what extent can it be said to be truly enactivist? Although this is the question that we aim to eventually address with our enactivist theory of personality, for

now we can begin establishing the plausibility of the theoretical link between RR and enactivism by considering the following three points. First, biological processes implement cost functions and RR describes the processes of cost function optimization as instantiated within a single dynamical economy. In this manner, RR can thus be extended into a *bio*economic theory of cognition (e.g., Vervaeke & Ferraro, 2013). Insofar as enactivism grounds mind in life, therefore, there is a possible theoretical link to be made between enactivism and RR.

Second, enactivist accounts of sense-making endeavor to explain cognition in terms of the dynamical or structural coupling patterns of autopoietic and adaptive systems (Di Paolo, 2005; Thompson, 2007). Building on Weber and Varela's (2002) seminal work in integrating autopoiesis with Hans Jonas' (1966) phenomenology of the organism, Di Paolo (2005) explains that what is characteristic of adaptive and autopoietic systems, contra other self-organizing systems, is that they bring forth their own normativity or "intrinsic teleology." This theoretical move seeks to provide a naturalistic framework for the intrinsic teleology of the organism by grounding it in the adaptive autopoietic organization of living beings. Accordingly, it is through adaptivity and autopoiesis that living beings presumably establish a frame of reference of their situation in accordance with their interests (e.g., self-continuation), consequently allowing enactivism to secure a non-reductionist stance toward subjectivity by introducing various phenomenological properties of embodied cognition into the ontological picture (e.g., an organismic "point of view," salience, concern, etc.).

And yet, an immanent problem for such accounts is likewise that of relevance: how does the cognitive system, embodied and enactive as it might be, "know" whether to be coupled with its environment in this way or in that way? How does it know how to realize what is relevant in its manners of coupling with the world without succumbing to combinatorially explosive strategies? Although enactivism construes cognition as a procedural and participatory (not primarily propositional) achievement of adaptive autopoietic systems, it still confronts and must circumvent the frame problem in its account of adaptivity—for to be adaptive means precisely to realize what is relevant (Hovhannisyan, 2021; Hovhannisyan & Dewey, 2017). Explaining adaptivity in a way that does not confront the frame problem is thus especially crucial if enactivism wishes not to fall prey to the same critiques as the very position (i.e., cognitivism) that it was originally meant to refute. In this way, we believe that RR as a bioeconomic theory of cognition can help to close this gap by simultaneously bridging up into the phenomenology of human sense-making while also bridging down into the dynamical processes through which adaptive autopoietic systems are able to realize their own immanent teleology.

We thus arrive at our third point, that a more substantial link can be established by interfacing RR with the phenomenological contributions of Merleau-Ponty, whose work, as we already know, constitutes the theoretical and philosophical backdrop of the enactive approach. Specifically, we think that RR can explain the underlying process dynamics of Merleau-Ponty's notion of "optimal grip," which construes the primary function of the mind as aimed at "skillful coping" with the environment (Bruineberg & Rietveld, 2014; Bruineberg et al., 2018; Dreyfus, 2014; Rietveld, 2008). To elaborate on what this means, consider the task of reading a book. Holding the book

too close or too far from one's face renders the text illegible and the task of reading, accordingly, impossible. Rather, one must hold the book at an optimal distance from one's face in order to effectively solve the problem of "reading," for only then does the text become legible. All cognitive problem solving, according to this account, therefore consists of acts of embodied interaction (skillful coping) involving processes of optimization like those involved in reading a book (e.g., optimizing the distance between the book and one's eyes). Thus, more than being inherently enactive, Merleau-Ponty's notion of optimal grip constitutes a conceptual metaphor that aptly reflects both the embodied nature of cognition as a process of "coming to grips" with worldly tasks as well as the dynamics of optimization that seem to be implicated in processes of this kind. In this light, we wish to advance the (even stronger) claim that RR can be readily integrated with Merleau-Ponty's notion in order to afford a more literal understanding of how optimal gripping is achieved by the embodied mind, thus elevating it from the status of mere conceptual metaphor to that of cognitive scientific theory. Our argument is that optimal gripping entails RR since optimization, which is a necessary condition for the optimal grip, technically requires and emerges from trade-off relationships, the basic building blocks of RR. As we see it, only a system that is doing opponent-processing between Efficiency and Resiliency can be said to be optimally gripping in a technical sense and thereby realizing relevance. Any dynamically situated cognitive system that is doing RR, in other words, is constitutively aimed at achieving an optimal grip over its world.

We believe this set of claims to be both theoretically plausible and experientially viable on several levels of analysis.[3] Despite the various theoretical advantages afforded by their framework of RR, Vervaeke et al. (2012) nevertheless admit that the cognitive economies of complex agents, such as human beings, likely entail other sets of constraints not yet articulated by their theory—though they do deem the system of constraints in their theory a plausible starting point for such an articulation. We agree with their self-assessment, for the most part. However, what we aim to show in the remainder of this article is that a more plausible articulation of the organization of the cognitive economies of human beings requires interfacing RR with CB5T. Whereas Vervaeke et al.'s (2012) original approach to RR is arguably more conceptually driven, ours can be said to be empirically motivated, for it entails a greater degree of engagement between RR, enactivism, and the empirical human sciences vis-à-vis personality theory—thereby exemplifying the kind of back-and-forth circulation that was said to be formative of early enactivist thought. Moreover, we find that such a circulation is

[3] Emerging neurological (Vervaeke & Ferraro, 2013) and biological (Crespi & Badcock 2008) evidence suggests the primacy of Efficiency- and Resiliency-based processing in human cognitive function. Psychologically, anxiety-related disorders serve as a relevant case example. Specifically, insofar as anxiety-related disorders are organized around rigid patterns of avoidance, they can be conceptualized as ineffective attempts based on Efficiency at eliminating sources of psychological entropy that happen to be otherwise irreducible (given the situation). Accordingly, exposure-based therapies, long predicated on the idea that voluntary (and guided) exposure to the sources of one's anxieties is necessary for curative change (Barlow 2014), can be conceptualized broadly as a process of facilitating Resiliency-based processing that is meant to replace the patient's rigidities with greater psychological flexibility.

also a necessary step toward articulating the immanent teleology of the autopoietically embodied human mind, a primary aim of this article and longstanding hope of the enactive approach. For on the one hand, the structural–functional organization of the cognitive economy as elucidated by RR, in its current form, is far too simplistic to be meaningfully related to the lived worlds of human beings, which are incomparably more layered and complex. Yet, on the other hand, the variables originally theorized about in RR were derived rationally (a priori), whereas those articulated by the Big Five have been derived empirically (a posteriori) and are thus by definition more plausible depictions of the de facto structures of human subjectivity. Thus, in seeking an integration of CB5T and RR, we can simultaneously (1) link personality with cognition, (2) thereby initiating a systematic dialogue between enactivist cognitive science and personality theory that is aimed at articulating the immanent teleology of the autopoietically embodied human mind, (3) all the while rendering RR more directly applicable to human lived experience by making it more deeply embodied.

5 Enactivist big five theory

A defining feature of the Big Five is that personality traits are organized hierarchically such that traits that are higher on the hierarchy subsume and are superordinate to traits lower on the hierarchy. The hierarchical structure of personality, though, emerged over time and as a consequence of careful statistical procedures. As various trait descriptors became grouped into broader categories via factor analysis the groupings themselves showed regular patterns that eventually yielded a complex covariance structure of traits existing at different levels and in relation to one another.

Initially, the Big Five were thought to be orthogonal and thus existing at the highest level of the hierarchy. However, it soon became clear that the Big Five are, "in fact, regularly intercorrelated such that there exist two higher-order traits, or *metatraits*, […] labeled *Stability* and *Plasticity*" (DeYoung 2006; DeYoung et al., 2002; Digman 1997 as cited in DeYoung, 2015, p. 36). In particular, meta-trait Stability accounts for the shared variance of traits Neuroticism, Agreeableness, and Conscientiousness, and meta-trait Plasticity, for the shared variance of traits Extraversion and Openness.

The literature as to the nature of what lies directly underneath the Big Five, on the other hand, has been somewhat mixed. For a long time, the prevailing point of view was that each trait subsumes six *facets*, comprising a total of 30 facets spread equally among the Big Five (McCrae & Costa, 2003). For example, standard personality assessment questionnaires used in the clinical context, such as the NEO-PI-R (Costa & McCrae, 1992), have assumed this model. A more recent model (i.e., the *Big-Five Aspects Scale* or *BFAS*), however, one that is advocated by DeYoung himself, has demonstrated that directly underneath the Big Five are actually 10 factors, or *aspects*, constituting an intermediary level between those of the traits and facets (DeYoung et al., 2007, 2013). According to this model, each trait has two aspects,

and each aspect counts as a higher-order factor in relation to the facets found lower in the hierarchy (DeYoung, 2015; DeYoung et al., 2007). DeYoung (2015) makes the following observation as to the statistical relations exhibited by the trait hierarchy with the aspects in mind:

> At each level of the hierarchy (below the highest [i.e., the meta-trait level]), some set of forces causes groups of traits to vary together in patterns described by the next higher level of the hierarchy, and some other set of forces causes each trait to vary independently of the others. In other words, all traits below the highest level of the hierarchy have both shared and unique valid variance. (p. 35)

What is interesting about the aspects is that, unlike the facets in the original model, they were derived empirically and thus pose a lower risk of researcher bias (DeYoung et al., 2007). As such, the ten aspects constitute a level of the personality hierarchy that is arguably more valid than that of the facets. What is even more interesting, however, is that the forces that "cause each trait to vary independently of the others" on one level, while also varying "together in patterns described by the next higher level of the hierarchy," seem to keenly resemble the structural–functional organization of the cognitive economy of RR in which distinct but complementary variables (e.g., special purpose and general purpose learning) are paired in trade-off relationships to optimize for some higher-order constraint (e.g., CS). We do not believe this semblance to be merely coincidental or superficial in its meaning. Indeed, we maintain that the statistical relationships among the traits in the hierarchy as described by DeYoung (2015) actually comprise empirical evidence of RR happening at various levels in the human personality. Particularly, we think that meta-traits Stability and Plasticity are functionally isomorphic to RR's higher-order constraints of Efficiency and Resiliency, respectively. Although not entirely identical in its overall structure to Vervaeke et al.'s (2012) original formulation of RR, the functioning of personality as it occurs within and between the various levels of the personality hierarchy nevertheless shares the same basic scheme. Seen in this light, we advance the following set of propositions regarding personality structure and function as a nested system of RR processes aimed at optimization:

(1) Structurally, each of the Big Five constitutes a constraint in the human cognitive economy
(2) Functionally, each of the Big Five emerges as a dynamical system from the opponent-processing of its respective aspect pairs which is optimizing for a given problem domain (of which there are a total of five)
(3) Structurally, the two meta-traits (Stability and Plasticity) constitute higher-order constraints in the human cognitive economy (Efficiency and Resiliency, respectively)
(4) Functionally, the two meta-traits emerge as dynamical systems from the opponent-processing of their respective trait-groupings (Neuroticism, Agreeableness,

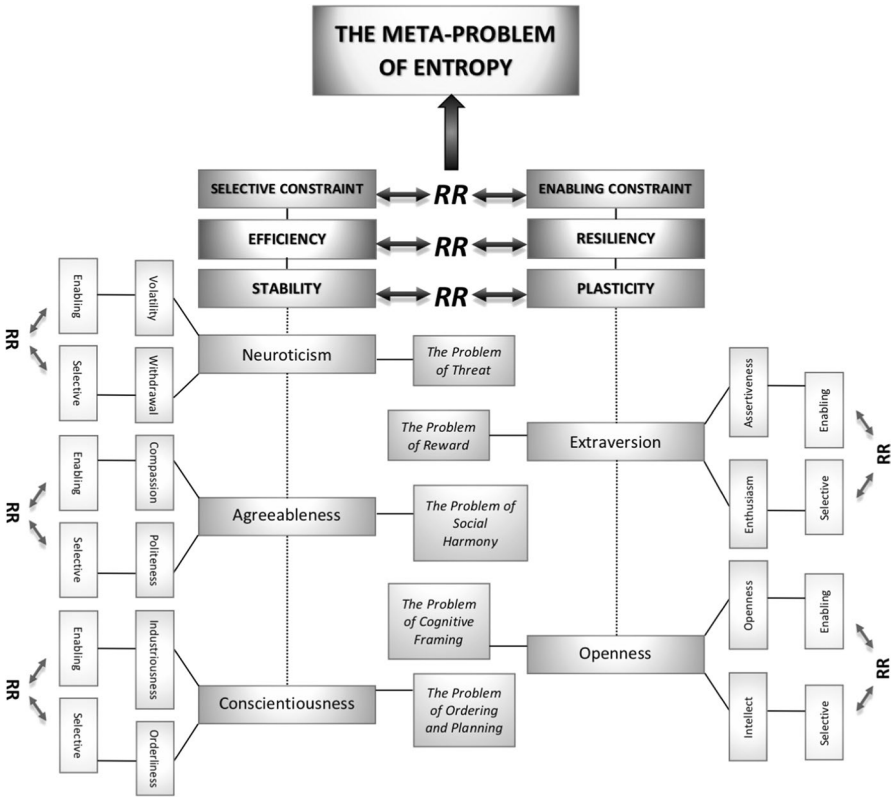

Fig. 1 EB5T Schematic of the Human Cognitive Economy

and Conscientiousness vs. Extraversion and Openness) which, on the whole, are optimizing for the meta-problem of entropy

(5) The meta-problem of entropy denotes the continuum of environmental conditions varying from highly entropic to highly static, with which any organism must have the capacity to contend by employing different and complementary styles of adaptation (i.e., Efficiency and Resiliency)

The theoretical argument we aim to advance in support of points (1)-(5) constitutes the conceptual backbone of our enactivist theory of personality, EB5T (see Fig. 1: EB5T Schematic of the Human Cognitive Economy). EB5T conceptualizes traits and their functions in a different manner from CB5T. Whereas CB5T conceptualizes the differential functionality of traits as parameters on the *cybernetic* system, thereby inviting cognitivism in through the front door, EB5T closes this door shut by conceptualizing the differential functionality of traits as parameters on the cognitive economy, namely as attractors influencing the dynamics of optimization of the various opponent pairs involved in the economy over longer timescales. Consequently, by grounding personality function in RR,

EB5T simultaneously evades the frame problem while also capturing in a more literal sense than CB5T the dynamical processes entailed in trait functioning.

The remaining task is to demonstrate how the human personality can be represented as a complex cognitive economy involving hierarchically organized opponent processes aimed at optimization across all levels of the personality. We must restrict our scope to just the meta-trait and trait levels, though we stipulate that all other levels could similarly be subjected to an opponent processing approach. Once completed, our analysis should make evident that traits are dispositional tendencies for how we come to optimally grip our distinctly human worlds.

6 EB5T and RR

RR structure and function and personality structure and function share a kind of organizational identity: both are patterns of situated dynamical activity but occurring on different timescales. On our account, personality refers to how a cognitive economy and all the opponent processes therein are organized and configured for optimal gripping. Traits are therefore cognitive systems nested within the broader personality hierarchy for skillfully coping with the world. Just as with RR, each trait acts as a constraint that (presumably) optimizes for functioning relative to a problem domain ("interactional problem," in Vervaeke et al.'s terms; 2012) for which the cognitive economy has adapted over evolutionary time. In this section, we articulate the various problems facing the cognitive economy at the trait level and explain how domain-specific coping is achieved by each trait through RR. We also discuss the main problem facing the cognitive economy at the meta-trait level (i.e., the meta-problem of entropy) and show how it is dealt with by tying all of our theoretical findings together into a coherent account of personality functioning.

6.1 Extraversion and the problem of reward

Trait Extraversion is a strong predictor of an individual's general level of sociability. High-scorers tend to be more cheerful, gregarious, outgoing, and amicable than their introverted counterparts (McCrae & Costa, 2003). However, the neurological evidence seems to suggest that the neural correlates of Extraversion are dopaminergically regulated and thus tied primarily to a person's reward systems more generally, implying that trait Extraversion is only indirectly associated with rewards found in social contexts (DeYoung, 2013, 2015).[4]

In CB5T, Extraversion is defined more broadly as "variation in parameters of the mechanisms designed to respond to rewards" (DeYoung, 2015, p. 42). Here, DeYoung (2015) argues that Assertiveness, one aspect of trait Extraversion, corresponds to feelings

[4] Given how the human environment is predominantly a social environment and that most sources of reward for human beings therefore emerge in social settings (e.g., receiving and seeking the positive attention of others at social gatherings), it should come as no surprise that trait Extraversion is nonetheless so strongly associated with a person's general level of sociability.

of *wanting* and is linked to the class of rewards called incentive rewards, which "indicate an increase in the probability of achieving a goal," while Enthusiasm, the second aspect, corresponds to feelings of *liking* and is linked to the class of rewards called consummatory rewards, which "represent the actual attainment of a goal" (p. 42) In a different article, DeYoung (2013) builds the case that dopaminergically regulated processes, such as those undergirding trait Extraversion (and trait Openness), are promoting exploration in light of perceived uncertainty. This hypothesis is predicated on the idea that the inherent value of uncertainty is neither positive nor negative, but irreducibly bi-valent (DeYoung, 2013; Hirsh et al., 2012; Peterson, 1999). To elaborate, although high-uncertainty settings present a greater risk of punishment (simply in virtue of their inherent indeterminacy) such that withdrawal via anxiety would be one adaptive way to respond, there is nevertheless inherent promise in engaging with uncertainty through exploration insofar as the unknown is not just where risks, but also opportunities for greater reward reside (Peterson, 1999). Trait Extraversion, on DeYoung's account, is therefore an adaptation to the incentive reward value of uncertainty and is predictive of how likely someone is to engage with uncertainty through such feelings as interest, curiosity, or excitement, rather than withdrawal by way of fear or anxiety (DeYoung, 2013, 2015). This explains why, for example, Extraverts are more naturally drawn to social events, like parties and gatherings, in which experiences that have an irreducible element of uncertainty are expected to occur (e.g., meeting new people, engaging in new conversation, going to new places, etc.).

We follow suit with DeYoung's characterization of the adaptive function of Extraversion but conceptualize it as a constraint in the human cognitive economy that is optimizing for the incentive reward value of uncertainty through processes of RR. Functionally, we think Extraversion emerges as a dynamical system that is doing opponent-processing between its two aspect pairs, Assertiveness and Enthusiasm. Here, Assertiveness can be conceptualized as an enabling constraint that generates possibilities of reward-seeking by incentivizing action, while Enthusiasm can be conceptualized as a selective constraint that reduces such possibilities for action by reinforcing reward-seeking behaviors that have yielded rewards in the past. Put differently, feelings of wanting generate possibilities for reward-seeking by making cues for reward more salient, while feelings of liking eliminate such possibilities by reinforcing affordances based on the strength of their perceived reward value. Thusly construed, Assertiveness and Enthusiasm are dynamically optimizing for the incentive reward value of uncertainty through recursive mutual feedback. Those who are high on both aspects will have their psychological life strongly organized around noticing, pursuing, and exploring the incentive reward value of uncertainty with a greater propensity for experiencing and being incentivized by such feelings as excitement, joy, curiosity, and enthusiasm. However, they will be at a disadvantage in such contexts that instead require a more cautious and calculated approach, in which the risk-to-reward ratio is higher, and self-regulation, rather than impulsivity, is rewarded. Conversely, though low Extraversion is a low-risk strategy, it places one at a disadvantage in contexts where exploration is rewarded and its associated risks are relatively low. It stands to reason, therefore, that trait Extraversion is optimally gripping for the incentive reward value of uncertainty, and the opponent-processing of Assertiveness and Enthusiasm realizes relevance when its engagement with what is potentially rewarding is occurring in a contextually sensitive manner.

6.2 Neuroticism and the problem of threat

If trait Extraversion is a dynamical system optimizing for the incentive reward value of uncertainty, Neuroticism is its counterpart: it is the constraint responsible for optimizing for security by managing the inherent *threat* in uncertainty (DeYoung, 2013, 2015; Hirsh et al., 2012). In CB5T, Neuroticism is defined as "the parameters that determine whether any increase in psychological entropy [uncertainty] triggers a defensive response" (DeYoung, 2013, p. 43). Phenomenologically, Neuroticism is therefore associated with the degree to which felt uncertainty is experienced as threatening or dangerous, and thus warranting, or even eliciting, a defensive response. High scorers in Neuroticism are more likely to experience feelings like fear or anxiety in the face of uncertainty and proceed with greater caution or defensiveness. Such individuals are therefore better adapted to situations that are high stakes, but less so when the context poses little to no risk, since their cautiousness or defensiveness causes them to lose out on meaningful opportunities. On the other hand, low measures of Neuroticism (i.e., Emotional Stability, the positive label for this trait) indicate a higher uncertainty threshold for one's defenses to be triggered. Phenomenologically, emotionally stable people are less likely to experience uncertainty as threatening, and less prone to experiencing and being perturbed by negative emotions, like anxiety, fear, anger, or depression. The downside of Emotional Stability is that it is more likely to result in false negatives because of a general hyposensitivity to threat. What is needed for optimization, therefore, is a process that can do RR with respect to the problem of threat in a contextually sensitive fashion.

DeYoung (2015) discusses the two aspects of Neuroticism, Volatility and Withdrawal, which on our account constitute the opponent functions that achieve RR in this domain. Volatility predicts one's tendency to engage in active defense and is linked to the fight-flight-freeze system (FFFS). Active defense involves such feelings as anger and panicked flight to "immediate threats or punishments where the only motivation is to escape or eliminate [perceived threats]" (p. 43). Thus, Volatility is an enabling constraint since it generates possibilities for action. Withdrawal, on the other hand, is linked to the behavioral inhibition system (BIS) and passive avoidance strategies, involving the "involuntary inhibition of approach toward a goal in response to increases in psychological entropy [uncertainty]" (DeYoung, 2015, p. 44). Thus, Withdrawal can be conceptualized as a selective constraint that reduces possibilities for action via inhibition. Trait Neuroticism accordingly emerges as a dynamical system optimizing for security through the opponent-processing of its two aspects, Volatility and Withdrawal. In other words, Neuroticism is optimally gripping for security when it is responding to possible threats either through active defense or passive avoidance in a contextually sensitive manner.

Interestingly, DeYoung (2015) hypothesizes that Neuroticism is regulated seratonergically and is opponent to Extraversion, which is regulated dopaminergically (p. 44). The claim that Neuroticism and Extraversion are opponent to one another has plausibility for two reasons. First, the functioning of the BIS and that of the behavioral activation system (BAS, a part of Extraversion) happen to be negatively correlated, which is how opponent processes tend to be related in general (DeYoung, 2015, pp. 43–44). And second, any adaptive system must do RR (at least partly)

by optimizing the trade-off between reward-seeking and threat-avoidance strategies, two basic processes necessary for survival. Traits Extraversion and Neuroticism, respectively, appear to be doing just that for a person. Individual differences in each of these traits therefore correspond to variation in one's general style of optimization with regards to the trade-off between threat-management and reward-seeking behaviors.

6.3 Openness and the problem of cognitive framing

Trait Openness has shared variance with trait Extraversion, which, on the neurological level, can be attributed to the fact that both are likely dopaminergically regulated (DeYoung, 2013). Whereas DeYoung refers to Extraversion as the trait responsible for behavioral exploration, he construes trait Openness as the functional basis for cognitive exploration (DeYoung, 2013, 2015). CB5T therefore defines trait Openness as the parameters on the mechanisms designed to respond to the incentive reward value of *information* (DeYoung, 2013, 2015). High scorers tend to be creative and outside of the box thinkers. They gravitate toward abstract ideas, enjoy intellectual discussions, and like novel and aesthetic experiences (McCrae & Costa, 2003). Low scorers, on the other hand, tend to be more concrete, conventional, and literal in their thinking. Openness is adaptive in contexts whereby the problems are ill-defined and creative engagement with information is rewarded (e.g., creating an art project or a novel theoretical framework). Conversely, Openness is less adaptive in contexts wherein problems are relatively well-defined and what is rewarded instead is the rote application of already known rules, principles, and knowledge (e.g., taking phone calls, booking appointments, etc.).

On our account, trait Openness is conceptualized as a constraint that functions to optimize the flexibility of one's cognitive framing. The two aspects of Openness that achieve this optimization are Intellect and openness (with a lowercase "o") (DeYoung, 2015). DeYoung (2015) explains that openness "has been linked to implicit learning," which is the "automatic detection of patterns in sensory experience" (p. 45). On the other hand, "the mechanisms of Intellect appear to be responsible for producing logical and causal knowledge about the world" (DeYoung, 2015, p. 45). We thus conceptualize openness as an enabling constraint that picks up on and makes salient correlational patterns through implicit learning, and Intellect as a selective constraint that extracts causal signal from correlational noise. Together, these two aspects dynamically optimize for the relevance of one's cognitive framing by varying the flexibility and fit of one's frames to the situation. Trait Openness is thus optimally gripping over cognitive possibilities for framing and realizes relevance when the opponent-processing of its two aspects is done in a contextually sensitive fashion.

6.4 Agreeableness and the problem of social harmony

Trait Agreeableness "represents the general tendency toward cooperation and altruism, as opposed to exploitation and lack of concern for others" (DeYoung, 2015,

p. 46). Agreeableness is thus a social trait and corresponds to the parameters on the processes for coordinating and prioritizing one's own needs and desires with those of others. Accordingly, too much agreeableness risks having one's own needs neglected whereas too much disagreeableness risks neglecting the needs of others.

The two aspects of Agreeableness that achieve optimization are Compassion and Politeness. Compassion involves bottom-up (involuntary and automatic) processes that promote prosocial behavior (such as feelings of compassion and empathy), whereas Politeness involves the top-down (voluntary and conscious) regulation of anti-social behaviors (such as anger and competition) (DeYoung, 2015, p. 46). We thus conceptualize Compassion as an enabling constraint of pro-social possibilities and Politeness as a selective constraint that makes pro-social possibilities more probable by eliminating anti-social possibilities. Through the opponent-processing of its two aspects, trait Agreeableness is therefore optimally gripping over possibilities for social harmony and realizes relevance in this domain when its processing is done in a contextually sensitive fashion.

6.5 Conscientiousness and the problem of ordering and planning

CB5T defines the final trait, Conscientiousness, as "variation in the mechanisms that allow people to follow rules and prioritize non-immediate goals" (DeYoung, 2015, p. 45). DeYoung observes that the interaction between Conscientiousness and traits Extraversion and Neuroticism is multifaceted and elaborates:

> The mechanisms of Conscientiousness are likely to have complex interactions with the reward-seeking and defensive motivational systems related to Extraversion and Neuroticism (Corr et al., 2013). In one situation, Conscientiousness might encourage suppressing an emotional reaction to a minor threat in order to pursue a non-immediate or abstract goal. In another situation, however, it might amplify attention to a very similar threat, if the latter was likely to interfere with the larger goal. Similarly, Conscientiousness should suppress reward-seeking that is a distraction from larger goals but encourage reward-seeking that furthers those goals. (DeYoung, 2015, p. 45)

DeYoung (2015) elaborates that Industriousness and Orderliness, the two aspects of Conscientiousness, "appear to reflect the distinction between prioritizing non-immediate goals and following rules" (p. 45). Together, he claims, the two aspects function to regulate motivational stability across time by either motivating work towards a long-term goal (Industriousness) or reducing distractibility (Orderliness). Too much conscientiousness risks neglecting all and any sources of immediate gratification whereas too little conscientiousness risks undermining the realization of longer-term goals of potentially greater value. Conscientiousness is optimal when the prioritization of long versus short-term goals is done in a contextually sensitive fashion.

On our account, therefore, Conscientiousness is a constraint in the human cognitive economy that is optimally gripping over the ordering and planning of goals existing on different timescales (Rueter et al., 2018). We conceptualize

Industriousness as an enabling constraint that functions to generate possibilities toward the realization of long-term goals, and Orderliness as a selective constraint that reduces potential distractors that may hinder the realization of long-term goals. Our thinking is in line with the empirical evidence according to which Industriousness is negatively related to Neuroticism and positively related to Extraversion (and hence motivated by greater sensitivity to potential rewards), whereas Orderliness is positively related to Neuroticism and negatively related to Extraversion (whereby rules are construed as protective strategies against defensive reactions to uncertainty, which may distract from the pursuit of goals) (Rueter et al., 2018).

6.6 Meta-traits stability and plasticity and the meta-problem of entropy

A fundamental problem facing cognitive agency is that the world is entropic whereas the frames by which the world is interpreted and rendered sufficiently predictable are static (DeYoung, 2015; Peterson, 1999; 2007; Thompson, 2007). This is a version of the frame problem discussed in Sect. 4. As such, when confronted with uncertainty, a decision must be made either to accommodate the world by changing one's frames or else to assimilate it into one's already existing frames. Neither strategy is totally effective, yet both are necessary and happen to be complementary to one another. In CB5T, it is the two meta-traits, Plasticity and Stability, which exist at the highest level of personality, that are functioning to meet each of these goals. Stability, which explains the shared variance of Conscientiousness, Agreeableness, and Neuroticism, is an adaptation to the inherently threatening value of uncertainty and functions to meet the goal of security by managing or reducing threat. On the other hand, Plasticity, which explains the shared variance of Openness and Extraversion, is an adaptation to the incentive reward value of uncertainty (behaviorally and cognitively) and functions to meet the goal of reward by promoting exploration.

On our account, meta-traits Stability and Plasticity are functionally isomorphic to the higher-order constraints of Efficiency and Resiliency in RR, respectively. What distinguishes the meta-traits and the original higher-order constraints is essentially the timescales at which they occur: whereas the meta-traits evidently pertain to trait-level dynamics (longer timescales), the higher-order constraints can be said to pertain to state-level dynamics (shorter timescales). This is to say, when considered from the timescale of traits, Efficiency corresponds to meta-trait Stability and Resiliency corresponds to meta-trait Plasticity. However, whereas in the original RR formulation, the higher-order constraints functioned by varying the possibilities for action available to the cognitive economy at any given point in time, at the level of traits we find it plausible that Stability and Plasticity are functioning to vary persons' dispositional tendencies or *styles* of optimally gripping their worlds. On the whole, we suspect that the opponent-processing of the two meta-traits (and the Big Five nested therein) is optimizing the person's *style of grip* over their problem landscape and their relationship with uncertainty in each of the five (trait-level) domains. Let us elaborate on what we mean by this.

Phenomenologically, we imagine that as a selective constraint, meta-trait Stability enacts a *style of skillful coping* that pushes for relatively enduring patterns

of person-situation fit. Stability is thus well-adapted mainly to contexts that are relatively static and unchanging and is also less metabolically costly. Moreover, it organizes patterns of enactment that are not only well-adapted to stable contexts but are likely to promote stability in contexts that are in flux (e.g., Conscientiousness may add structure to an unpredictable situation, Agreeableness may promote social harmony in a time of conflict, Emotional stability may reduce distractibility from reward-seeking). On the other hand, we imagine that meta-trait Plasticity enacts a complementary style of skillful coping that is pushing, instead, for greater *evolvability* of the person-situation fit. Specifically, we conceptualize Plasticity as an enabling constraint that promotes cross-contextual flexibility (e.g., Extraversion promotes opportunism through behavioral exploration, Openness promotes cognitive flexibility through complexity in processing). Thus, Plasticity functions by organizing patterns of enactment that are not only well-adapted to dynamic contexts but might also promote change in contexts that are relatively static (e.g., making jokes or creating novel scientific theories, both fulfil this function in their respective domains). Yet it is for this reason that Plasticity necessarily places the system into a more metabolically costly state, thereby risking the system's Resilience (at the state level) if maintained as a long-term strategy. In light of these claims, we think that the autopoietically embodied human mind is doing RR at the highest level of analysis when it is optimizing the trade-off between Stability and Plasticity in a contextually sensitive fashion. A well-functioning personality is thus one in which the Big Five are promoting both Stability and Plasticity optimally (DeYoung, 2015, p. 46). Conversely, a dysfunctional personality is one that has fallen prey either to Rigidity (the pathological pole of Stability) or to Instablity (the pathological pole of Plasticity) and is no longer optimally gripping its problem landscape (DeYoung, 2015).

7 EB5T Beyond cognition and personality

Evolutionary forces seem to favor Efficiency over Resiliency-based strategies for coping, since metabolic resources are limited yet vital to our continued survival. How is it that human beings exemplify such a degree of Plasticity, then, despite all biological odds? It is likely that the optimization of this trade-off is not merely taking place at the level of individual personality functioning (between meta-traits Stability and Plasticity *within* an individual), but also between individuals, at the level of the population or group. Not all individuals can exhibit high levels of Plasticity in their functioning as that would be too metabolically costly and disadvantageous. But if just enough individuals exhibit relatively high levels of Plasticity, while the rest tend toward Stability, then the group as a whole will achieve a degree of resilience in its distributed processing that might not be true of its particular individuals, thereby attaining optimality in the trade-off between Stability and Plasticity. In other words, the group as a whole becomes a "master of all trades." Hence, we theorize that personality type variation within populations of individuals is evolution's way of optimizing the fundamental trade-off between Plasticity and Stability, namely by securing a degree of Plasticity (and Resiliency) at the level of groups through means of distributed cognition, while maintaining individuals' metabolic costs within an

optimal range of expenditure by pushing for Stability (and Efficiency) at the individual level.

We would predict that due to this evolutionary bias toward Stability, there would be a greater preponderance in individuals toward Rigidity (and various forms of psychopathology that stem from Rigidity), but that the negative effects of this bias would likely be ameliorated through social means by the Resilience introduced through the distributed cognition of the group (e.g., through the development and inheritance of cultural norms, rituals, social practices, healing practices like psychotherapy, etc.). To this end, we hypothesize that culture comes in to fill this void by off-loading the problem of framing from the individual to the group, so as to ease the cognitive load placed on each particular individual and reduce metabolic expenditure. This is to say, culture functions as a mechanism for solving the frame problem and evading combinatorial explosion for the individual through means of distributed cognition: it is one of the forces that provides individuals with "ready-made" frames for optimally gripping the world so that they are not overwhelmed with having to do all of the gripping individually and from the ground up. Here, we think that Ramstead et al.'s (2016) work at the intersection of cultural anthropology, enactivist cognitive science, and predictive processing, is particularly germane for conceptualizing how Stability and Plasticity-based dynamics unfold at the level of groups and feed into the individual cognition of their members. Although this is not a point of connection we can explore or develop in our paper, we see it as a worthwhile direction for EB5T to move toward in the future when broaching questions of cultural diversity from a cognitive science lens as well as exploring the continuity between (embodied) cognition, personality, and culture, more broadly.

8 Conclusion

The guiding theme of this article has been that a theoretically mature enactivism is bound to be humanistic in its articulation, and only by becoming more humanistic can enactivism more fully embody the non-reductionist spirit that lay at its foundation. We have sought to bring forth such an articulation of the embodied-enactive mind by developing an enactivist theory of personality, EB5T, through a massive synthesis of core concepts from personality theory (CB5T), cognitive science (RR), and phenomenological philosophy (optimal gripping). In short, EB5T explains the immanent teleology of the autopoietically embodied human mind as a kind of full-scale optimal gripping process that is achieved along the five major dimensions of personality, which are said to describe universal human structures. We have sought to demonstrate that personality traits are dispositional tendencies for how we come to optimally grip our distinctly human worlds and that individual differences in personality are reflective of stylistic differences in optimal gripping tendencies between human beings. To the degree that these structures (the traits and meta-traits) are indeed universal, we therefore believe that our theory offers a viable path forward in advancing enactive cognitive science beyond the life of a cell and into the mind of a person, a longstanding hope and ambition held by proponents of the enactive approach. Moreover, through our

discussion of EB5T and distributed cognition, we have introduced a novel angle from which to consider the adaptive functionality of traits, namely that of group-level optimization, thereby establishing a direct continuity between biology, cognition, personality, and culture. This extended continuity hypothesis forms an area of research that is fertile with the possibility of interdisciplinary engagement, though we are especially eager to see such engagement from proponents of CB5T, like DeYoung, whose contributions to a cognitive scientific understanding of personality function have thus far been ground-breaking and deeply informative of our own work.

Before concluding, we think it important to also discuss the relevance of EB5T for (1) personality research and (2) clinical psychopathology and psychological assessment practice. With regards to the former, we believe that EB5T—particularly in virtue of its phenomenological foundations—affords a technical means of reconceptualizing personality traits from properties that are internal to oneself to ways in which one *tends* to participate in and experience one's world. As a means of capturing the world-involving nature of personality, therefore, we think that it might be appropriate to shift from the language of describing personality traits in terms of structures and functions to a language that describes them as *styles of world-enactment* (Fischer, 1994). Accordingly, such a reconceptualization bears important implications regarding the nature of person-situation interaction, whereby person and situation are no longer to be understood as apart from each other, but as mutually inseparable parts of the same dynamical whole.

With regards to psychopathology, EB5T advances a criterion of psychological function and dysfunction in optimal gripping terms. On the one hand, EB5T agrees with CB5T that trait measures on their own do not necessarily indicate psychopathology (DeYoung & Krueger, 2018; DeYoung & Weisberg, 2018). Yet EB5T departs from CB5T by endorsing an understanding of traits as fundamentally situated and world-involving. The implications of adopting a situated understanding of trait function are twofold. First, that as predispositions for skillful coping, all traits are optimal only in some contexts but not in others. And second, that dysfunction—or suboptimal gripping—ensues when situational demands exceed a trait's (finite) capacity to adapt. For example, we can imagine how opportunity-seeking behaviors that are motivated by Extraversion can be adaptive in low-risk contexts (e.g., partying with friends) but can become problematic when risk levels are significantly higher (e.g., partying with friends amidst the COVID pandemic) and what is rewarded instead is prudence (e.g., remaining at home and in small groups). Since traits tend to be stable across contexts (McCrae & Costa, 2003), whether they act as functional (optimal) or dysfunctional (suboptimal) depends on how well they fit with and are able to meet the unique demands of the situation. Using tests that are derived from the Big Five (e.g., NEO-PI-R) *diagnostically* thus requires having working knowledge of the client's context and personality traits in order to determine if there is a mismatch of the two.

On the basis of these theoretical principles, we advance four testable hypotheses with regards to possible diagnostic uses of the Big Five for the assessment of psychopathology:

- **H1**: Mismatches in what a situation calls for (e.g., prudence) and how a trait is designed to function (e.g., opportunistic behaviors motivated by high Extraversion) will lead to *dysfunction* and cause psychological distress (e.g., the "Lonely Extravert")
- **H2**: Individuals will form cognitions (e.g., beliefs and judgments) about their experiences of distress (e.g., "I feel alone in this pandemic" or "no one wants to spend time with me")
- **H3**: Individuals' cognitions will relate thematically to Big Five traits that are entailed in the underlying mismatch (e.g., Extraverts' concerns around feeling lonely will naturally reflect a frustration of their need for social engagement), thereby implicating the potential source(s) of dysfunction; and
- **H4**: Sources of trait-situation mismatch can be identified in patients' descriptions of what is bothering them

In light of these hypotheses, how might an EB5T-inspired approach to assessment look in practice? Let us consider one such possibility before concluding. After norming interviewees' scores on the NEO and generating their profiles, assessors could begin by adopting a collaborative approach and enlisting interviewees in the sense-making process to reach meaningful understandings of the findings together (Fischer, 1994; Finn et al., 2012). Contextualizing and thematizing interviewees' personality profiles with real examples drawn from their lived experience can be useful for drawing out the assessment's full meaning and impact during the interview process (Fischer, 1994; Finn et al., 2012). Once assessors have discussed the scores and their possible meanings with participants, they could then pose such questions as, "Where do you recognize yourself in these scores?", "How might these scores reflect your personal strengths or growing edges?", or "When has this trait benefited you and when has it gotten in the way?" Raising these sorts of questions will not only invite interviewees to consider their personality functioning as it is experienced in their lives, but also to realize the contextual nature of their individual strengths and weaknesses, and, with it, the possibility of psychological change. Assessors could then follow up by having interviewees imagine what they would do differently if they were to capitalize on their particular strengths while working to improve their weaknesses, thus hitting on the most central tenet of our theory: optimization.

Evaluating the plausibility of our theoretical contribution will ultimately require exploring and testing ways that EB5T can be used to (systematically) inform assessment and intervention procedures in clinical and non-clinical settings. Our illustration in the previous paragraph serves only as a hypothetical example of what is otherwise an empirical question. The broad aim of future studies should be to develop methods of assessment that are directly geared at facilitating optimization, or optimal gripping, in line with the various principles outlined by our theory (e.g., opponent-processing, trade-off relationships, etc.).

As an alternative to CB5T, it is also important to discuss possible points of convergence and divergence between EB5T and its predecessor. One possible critique of our enactivist reformulation of CB5T is that CB5T is already inherently enactivist as per Safron and DeYoung's (2020) recent work in integrating concepts from CB5T with the free energy principle (FEP) and

active inference (AI) framework—both of which have been thoroughly explored and situated in an enactivist context (e.g., Engel et al., 2015; Gallagher, 2017; Kirchhoff, 2018).[5] Put differently, the question stands as to what degree CB5T is already enactivist and whether EB5T is able to contribute anything truly unique or significant to the picture that can differentiate it from CB5T. We offer the following remarks in response to this set of concerns. First, although both frameworks employ concepts from dynamical systems theory in their explanatory pursuits (e.g., recursive processes of self-organization), we believe that EB5T currently has a theoretical edge over CB5T at least insofar as EB5T's account of cognition (i.e., RR as optimal grip) explicitly circumvents one of the most notorious and central problems in all of cognitive science: the frame problem. We admit, however, that we do not see this as a permanent limitation for CB5T. Indeed, we believe that there is significant potential to integrate RR theory with an enactivist reading of the FEP-AI framework that was recently used to ground CB5T (Safron & DeYoung, 2020), in which RR could be explained more specifically as a function of precision-weighting differences in how cognitive economies optimize various trade-offs among their hierarchically organized predictive processes. Such an integration would at least in principle inoculate CB5T against the frame problem through its grounding in RR while demonstrating further convergence between CB5T and EB5T toward an understanding of personality trait function in predictive processing terms.

In addition to this point, Safron and DeYoung (2020) have noted potential correspondences between CB5T's meta-traits Stability and Plasticity and FEP-AI's respective optimization for the extrinsic value of satisfying prior preferences and the intrinsic value of epistemic exploration. This is highly consistent with our suggested correspondence between Stability and Plasticity and the higher-order constraints of Efficiency and Resiliency in the RR framework, which grounds EB5T. Such correspondences between independently developed concepts in EB5T, RR, CB5T, and FEP-AI are non-trivial and provide much stronger grounds for confidence in the concepts utilized in these frameworks than any one of these frameworks is able to establish in isolation. Such a degree of convergence not only adds plausibility to EB5T, RR, CB5T, and FEP-AI but also provides further support for potential compatibility between and integration of EB5T and RR with what is arguably the first unified paradigm for formal modeling in the mind and life sciences: FEP-AI. Moreover, it affords points of further dialogue between proponents of CB5T and those working in the enactivist paradigm—for example, by affording a means of conceptualizing optimal grips, which are a fundamentally enactivist concept, in predictive processing terms.

Despite these various points of convergence between EB5T (and RR) and CB5T (and FEP-AI)—and especially CB5T's extant enactivist leanings—we believe that what distinguishes EB5T from CB5T *philosophically* is the former's phenomenological conceptualization of personality traits and the applicability of such

[5] Despite these apparent alliances between FEP-AI and enactivist thought, some have recently contended whether the implicit assumptions of the free energy approach (FEA) are indeed compatible with many of enactivism's own core concepts (e.g., adaptivity; see Di Paolo, Thompson, & Beer 2021). It is beyond the scope of our paper to review these debates and to provide definitive solutions. Our purpose in making mention of this is simply to note that the degree of compatibility between enactivism and FEA—and therefore of CB5T, which makes use of FEA—is not settled but ongoing.

understandings to clinical practice. Whereas for CB5T traits reside *in* a person, EB5T makes possible a different kind of hermeneutic engagement with human personality altogether. According to EB5T and its world-involving conception of traits, if we wish to better understand who someone is, we must look beyond their personality measures and turn to their world. As we have argued, individual differences in personality do not merely reflect differences in optimal gripping tendencies but also differences in *styles of world-enactment*—a structural feature of personality that is *irreducible* and must be taken into account by clinicians wishing to helpfully intervene in persons' lives. Conversely, CB5T is grounded in a family of traditions whose philosophical spirit is not fundamentally enactivist precisely because it lacks a phenomenological foundation—a foundation without which it cannot (in principle) claim to be non-reductionist in the exact sense that enactivism (at least the variety with which we have aligned ourselves here) seeks to be.[6] Of course, it is possible to derive phenomenological understandings of trait differences on the basis of CB5T in order to bridge this gap and, accordingly, to afford the kinds of clinical understandings made possible by EB5T. But the moment one does this, one has already departed from the province of CB5T and entered that of EB5T, for one is now conceiving of personality in the same radical (non-reductionist) spirit that enactivism had originally adopted in relation to human experience. On our view, it is precisely because of its phenomenological foundation that EB5T is properly enactivist and CB5T is not (despite its various enactivist sub-elements, e.g., FEP-AI). And it is precisely the gap between the human mind's organizational and phenomenological properties—a gap that is of central importance to the enactivist, not cybernetics, paradigm—that we have sought to bridge through our theoretical contribution. In this light, we strongly believe that our article exemplifies the kind of interdisciplinary dialogue that was first set in motion by Varela et al. (1991) some thirty years ago, and we hope that it will act as a roadmap for those wishing to engage with human experience and personality in a similar light.

References

Allen, T. A., & DeYoung, C. G. (2016). Personality neuroscience and the five-factor model. In Thomas A. Widiger (Ed.), *The Oxford Handbook of the Five Factor Model of Personality*, https://doi.org/10.1093/oxfordhb/9780199352487.013.26

Barlow, D. H. (Ed.). (2014). *Clinical handbook of psychological disorders: A step-by-step treatment manual*. The Guilford Press.

Bockelman, P., Reinerman-Jones, L., & Gallagher, S. (2013). Methodological lessons in neurophenomenology: Review of a baseline study and recommendations for research approaches. *Frontiers in Human Neuroscience, 7*, 608.

Bourgine, P., & Stewart, J. (2004). Autopoiesis and cognition. *Artificial Life, 10*, 327–345. https://doi.org/10.1162/1064546041255557

[6] In this article, we align ourselves with Thompson's (2004; 2007) understanding of the enactive approach as a fundamentally phenomenology-driven approach to cognitive science whereby to omit the phenomenological dimension from our understandings of mental life would be to undermine the very non-reductionism that enactivism seeks to uphold.

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese, 195*, 2417–2444.
Bruineberg, J. & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience, 8*, https://doi.org/10.3389/fnhum.2014.00599
Cervone, D., & Mischel, W. (Eds.). (2002). *Advances in personality science*. The Guilford Press.
Cherniak, C. (1986). *Minimal rationality*. MIT Press.
Corr, P. J., DeYoung, C. G., & McNaughton, N. (2013). Motivation and personality: A neuropsychological perspective. *Social and Personality Psychology Compass, 7*,158–175.
Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment, 4*, 5–13.
Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences, 31*, 241–340.
Dennett, D. (1987). Cognitive wheels: The frame problem of AI. In Z. W. Pylyshyn (Ed.), *The robot's dilemma: The frame problem in artificial intelligence* (pp. 41–64). Ablex.
DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91,* 1138–1151.
DeYoung, C. G. (2010). Personality neuroscience and the biology of traits. *Social and Personality Psychology Compass, 4*(12), 1165–1180. https://doi.org/10.1111/j.1751-9004.2010.00327.x
DeYoung, C. G. (2013). The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Frontiers in Human Neuroscience,* 7. https://doi.org/10.3389/fnhum.2013.00762
DeYoung, C. G. (2015). Cybernetic Big Five Theory. *Journal of Research in Personality, 56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004
DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*, 533–552.
DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, R. J. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science, 21*(6), 820–828. https://doi.org/10.1177/0956797610370159
DeYoung, C. G., & Krueger, R. F. (2018). Understanding Psychopathology: Cybernetics and Psychology on the Boundary between Order and Chaos. *Psychological Inquiry, 29*(3), 165–174. https://doi.org/10.1080/1047840X.2018.1513690
DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between Facets and Domains: 10 Aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880.
DeYoung, C. G., & Weisberg, Y. J. (2018). Cybernetic Approaches to Personality and Social Behavior. In Deaux, K. & Snyder, M. (Eds.) *The Oxford Handbook of Personality and Social Psychology* (2nd ed.), https://doi.org/10.1093/oxfordhb/9780190224837.013.25
DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the Aspects of the Big Five, the Interpersonal Circumplex, and Trait Affiliation. *Journal of Personality*. https://doi.org/10.1111/jopy.12020
Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4*(4), 429–452.
Di Paolo, E., Thompson, E., & Beer, R. D. (2021). Laying down a forking path: Incompatibilities between enaction and the free energy principle. https://doi.org/10.31234/osf.io/d9v8f
Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256
Donnellan, M. B., Conger, R. D., & Bryant, M. C. (2004). *Journal of Research in Personality, 38*, 481–504. https://doi.org/10.1016/j.jrp.2004.01.001
Dreyfus, H. (2014). *Skillful coping: Essays on the phenomenology of everyday perception and action*. Oxford University Press.
Engel, A. K., Friston, K. J., & Kragic, D. (2015). *The pragmatic turn: Toward action-oriented views in cognitive science*. MIT Press.
Fischer, C. (1994). *Individualizing psychological assessment*. Lawrence Erlbaum Associates Inc.
Finn, S. E., Fischer, C. T., & Handler, L (2012) *Therapeutic assessment: A casebook and guide* (Eds.). Hoboken, New Jersey: John Wiley & Sons, Inc
Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality, 56*, 82–92. https://doi.org/10.1016/j.jrp.2014.10.009

Froese, T. (2011). From second-order cybernetics to enactive cognitive science: Varela's turn from epistemology to phenomenology. *Systems Research and Behavioral Science, 28*(6), 631–645.
Froese, T., & Di Paolo, E. (2009). Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences, 8*, 439–463. https://doi.org/10.1007/s11097-009-9140-8
Fuchs, T. (2018). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford University Press.
Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press.
Gordon, S. (Ed.). (2013). Neurophenomenology and its applications to psychology. Springer. https://doi.org/10.1007/978-1-4614-7239-1
Hirsh, J. B., Mar, R. A., & Peterson, J. B. (2012). Psychological entropy: A framework for understanding uncertainty-related anxiety. *Psychological Review, 119*(2), 304–320.
Hovhannisyan, G. (2018). Humanistic cognitive science. *The Humanistic Psychologist, 46*(1), 30–52.
Hovhannisyan, G. (2021). Specifying the conditions for a theory of teleology in cognitive science. *Journal of Theoretical and Philosophical Psychology,* Advance online publication. https://doi.org/10.1037/teo0000184
Hovhannisyan, G., & Dewey, C. (2017). Natural & normative dynamical coupling. *Cognitive Systems Research, 43*, 128–139.
Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*(6), 869–879. https://doi.org/10.1037//0021-9010.85.6.869
Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy. Carr, D. (trans.)*. Northwestern University Press.
John, O. P., Robins, R. W., & Pervin, L. A. (2008). *Handbook of personality: Theory and research* (3rd ed.). The Guilford Press.
Jonas, H. (1966). *The Phenomenon of Life: Toward a Philosophical Biology.* Harper and Row.
Juarrero, A. (1999). *Dynamics in action: Intentional behavior as a complex system*. MIT Press.
Kirchhoff, M. D. (2018). Autopoiesis, free energy, and the life-mind continuity thesis. *Synthese, 195*(6), 2519–2540.
Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: In support of a strong life-mind continuity thesis. *Entropy 19*(4), https://doi.org/10.3390/e19040169
Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences, 51*, 472–477. https://doi.org/10.1016/j.paid.2011.04.019
Lutz, A., & Thompson, E. (2003). Neurophenomenology- Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies, 10*(9–10), 31–52.
Major, D. A., Turner, J. E., & Fletcher, T. D. (2006). Linking proactive personality and the big five to motivation to learn and development activity. *Journal of Applied Psychology, 91*(4), 927–935. https://doi.org/10.1037/0021-9010.91.4.927
McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). The Guilford Press.
McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2010). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*(1), 28–50. https://doi.org/10.1177/1088868310366253
Merleau-Ponty, M. (1942/1963). *The Structure of Behavior*. Fisher, A. (trans.), Duquesne University Press
Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Hill and Wang.
O'Connor, M. C., & Paunonen, S. V. (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*, 971–990. https://doi.org/10.1016/j.paid.2007.03.017
O'Regan, J., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–973. https://doi.org/10.1017/S0140525X01000115
Olivares, F. A., Vargas, E., Fuentes, C., Martinez-Pernia, D., & Canales-Johnson, A. (2015). Neurophenomenology revisited: Second-person methods for the study of human consciousness. *Frontiers in Psychology, 6*, 673.
Peterson, J. B. (1999). *Maps of meaning: The architecture of belief*. Routledge.
Petitmengin, C., Navarro, V., & Le Van Quyen, M. (2007). Anticipating seizure: Pre-reflective experience at the center of neuro-phenomenology. *Consciousness and Cognition, 16*, 746–764.

Ramstead, M. J. D., Vessière, S. P. L., & Kirmayer, L. J. (2016). Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology, 7,* https://doi.org/10.3389/fpsyg.2016.01090

Rietveld, E. (2008). The skillful body as a concernful system of possible actions: Phenomena and neurodynamics. *Theory and Psychology, 18*(3), 341–363. https://doi.org/10.1177/0959354308089789

Robbins, B. (2013). Enactive cognition and the neurophenomenology of emotion. In S. Gordon (Ed.), *Neurophenomenology and its application to psychology (pp. 1–24).* Springer. https://doi.org/10.1007/978-1-4614-7239-1_1

Robbins, B., & Gordon, S. (2015). Humanistic neuropsychology: The implications of neurophenomenology for psychology. In K. Schneider, J. Pierson, & J. Bugental (Eds.), *The handbook of humanistic psychology: Theory, research, and practice (pp. 196–212).* Sage. https://doi.org/10.4135/9781483387864.n15

Rueter, A. R., Abram, S. V., MacDonald, A. W., III., Rustichini, A., & DeYoung, C. G. (2018). The goal priority network as a neural substrate of conscientiousness. *Human Brain Mapping, 39*(9), 3574–3585. https://doi.org/10.1002/hbm.24195

Safron, A., & DeYoung, C. G. (2020). Integrating Cybernetic Big Five Theory with the Free Energy Principle: A new strategy for modeling personalities as complex systems. https://doi.org/10.31234/osf.io/653wp

Sundararajan, L., & Ho, W.-T. (2020). Order against the tide of entropy: Self-distancing in expressive writing explained. *The Humanistic Psychologist, 48*(2), 183–201. https://doi.org/10.1037/hum0000146

Thompson, E. (2004). Life and mind: From autopoiesis to neurophenomenology, a tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences, 3*, 381–398. https://doi.org/10.1023/B:PHEN.0000048936.73339.dd

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences, 4*, 407–427. https://doi.org/10.1007/s11097-005-9003-x

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Belknap of Harvard UP.

Thompson, E. (2011). Précis of mind in life: Biology, phenomenology, and the sciences of mind. *Journal of Consciousness Studies, 18*(5–6), 10–22.

Varela, F. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies, 3*, 330–349.

Varela, F. J. (1999). The specious present: A neurophenomenology of time consciousness. In J. Petitot, F. J. Varela, B. Pachoud, & J. M. Roy (Eds.), *Writing science. Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science (p. 266–314).* Stanford University Press.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT.

Vervaeke, J., & Ferraro, L. (2013). Relevance realization and the neurodynamics and neuroconnectivity of general intelligence. In I. Harvey, et al. (Eds.), *Smart data: Privacy meets evolutionary robotics* (pp. 57–68). Springer. https://doi.org/10.1007/978-1-4614-6409-9_6

Vervaeke, J., Lillicrap, T., & Richards, B. (2012). relevance realization and the emerging framework in cognitive science. *Journal of Logic and Computation, 22*(1), 79–99.

Vörös, S. (2020). Mind embodied, mind bodified: Merleau-Ponty and the enactive turn in mind sciences. *Phenomenological Studies, 4*, 91–117.

Weber, A., & Varela, F. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences, 1,* 97–125.

Widiger, T. A., & Costa, P. T. (2013). *Personality disorders and the five-factor model of personality* (3rd ed.). American Psychological Association.

Zhao, H., & Seibert, S. E. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology, 91*(2), 259–271. https://doi.org/10.1037/0021-9010.91.2.259