



BOOK REVIEW:  
*Self-Reflective Fiction and 4E Cognition*

*Essi Varis*

Polvinen, Merja. *Self-Reflective Fiction and 4E Cognition*. Routledge, 2023.  
ISBN 9781032263748.

Full disclosure: I am fortunate enough to have known cognitive narratologist Merja Polvinen for many years now, and during those years, I have developed a certain heuristic. Whenever a scholarly text I am reading discusses the simultaneous life-likeness and artificiality of fiction, I scribble in the margin: “the Merja Argument,” sometimes accompanied by a little smiley face. Other people might signpost such passages with more general terms, such as “the paradox of fiction,” but one has to be careful there, because – as *Self-Reflective Fiction and 4E Cognition* details – there is actually a crucial difference between the two viewpoints.

“The paradox of fiction” denotes the human tendency to care about fictional happenings as if they were real, while knowing full well that they are not. The paradox thus rests on the widely assumed ontological division between the immersive “belief” in the seeming reality of storyworlds and the more “detached” appreciation of the carefully crafted textual artifice. In narrative theory, these two levels of reading have typically been viewed like water and oil: they cannot mix in the vessel of consciousness. Thus, the best the reader can do is to swing their attention back and forth between the mimetic illusion and the textual-material design. “The Merja Argument,” by contrast, maintains that there is no ontological divide and, hence, no real paradox between the storyworld and the story design. Rather, reading is a form of “lucid self-delusion” (155), where the reading mind and the readable text open each other to the coinciding possibilities of mimesis and artifice simultaneously – inseparably. What is more, Polvinen goes on to argue that literary narratives constitute particular “cognitive environments” for fostering such layered effects of fictionality and self-reflection. Thus, interacting with them can potentially

give rise to entirely new “constellations” of thought (161–62). This mind-expanding potential of narratives is further underlined by Polvinen’s engagements with several carefully selected works of speculative fiction – in other words, works of fiction that expressly encourage their readership to reach somewhere beyond everyday realities.

These attractive new starting points for understanding literary imagination derive from a bundle of widely recognized narratological theories on the one hand, and from the recent wave of so-called “4E” cognitive theories on the other hand. The 4E paradigm is notoriously difficult to explain in any concise manner, but the first chapter of the book defines “cognition” as skillful action that is always deeply Embodied as well as Extended and Embedded in different kinds of environments. In explaining the fourth “E” – Enaction – Polvinen relies heavily on J. J. Gibson’s concept of “affordances,” “the action potentials offered by an environment to a particular cognizer” (5). This is the main move that allows Polvinen to redefine fictions as unique cognitive environments affording unique cognitive actions. The four subsequent chapters go on to explore the spectrum of the metacognitive qualities that differentiate these fiction-induced cognitive actions from everyday cognition: the metaphorical mode of understanding, the spatio-temporal structures of narratives, and fictionality.

Before peeking into the main arguments of the analysis chapters, however, a word of caution may be in order: *Self-Reflective Fiction and 4E Cognition* is one of those books that are far heavier than their size would suggest. Anyone who has read the first chapter of the book – or the preceding three paragraphs of this review for that matter – surely realizes that *Self-Reflective Fiction and 4E Cognition* traverses fairly abstract levels of metafiction and metacognition. It does not satisfy itself with merely analyzing metaphorical and spatio-temporal structures of narratives; it also critiques the spatial metaphors that proliferate in narratology. Polvinen suggests, for instance, that narratologists’ habit of thinking about what is “inside” and “outside” of storyworlds may be one reason for the unnecessarily ontological divide perceived between the mimetic and the artificial aspects of fiction. In addition, Polvinen’s undeniable expertise in narrative theory means that her argumentation is quite brief and to-the-point. There is no idle rambling in this book, which other cognitive narratologists are sure to appreciate. For others, however, these same qualities might make the discussion rather dense and difficult to approach. One can only hope that the combination of these fairly high levels of abstraction – both in terms of the topic and in terms of the writing style – will not deter potential readers, because *Self-Reflective Fiction and 4E Cognition* would offer valuable takeaways to narratologists, SFF researchers, and cognitive scientists alike.

The narratologists can look forward to interesting terminological discoveries. Polvinen suggests, for instance, that the “double vision” of fiction as both life-like and artificial as well as the analogous “seeing-as” prompted by metaphorical language can both be captured in A. S. Byatt’s notion of “lamination” – “a word she uses to describe the simultaneous presence of layers of experience” (27). In the chapter concerning temporal structures, Polvinen applies the idea of “vection,” “a partial experience of self-movement” created by eye movement, to the cognitive movement across narratological frames (98–99). Moreover, in the chapter concerning spatial structures, Polvinen’s own

coinage, “making space,” demonstrates the two-way interactionality of affordances: it refers both to the readers’ ability “to imagine the text in spatial terms” and to the text’s way of “becoming a space for them to inhabit” (76–77).

The SFF researchers, for their part, are sure to enjoy the discussions Polvinen orchestrates between her theoretical musings and the chosen literary samples. As I have already mentioned, the book’s self-reflective corpus consists not only of traditional metafiction – works by Byatt, Ali Smith, John Barth, and David Eggers – but also of speculative fiction: Catherynne M. Valente’s “Silently and Very Fast” (2011), China Miéville’s *The City & the City* (2009), Ted Chiang’s “Story of Your Life” (1998), and Christopher Priest’s *The Prestige* (1995). Seeing that the book’s overarching mission is to explore how thinking with fiction differs from real-life cognition, focusing on overtly artificial postmodern narratives and overtly imaginary speculative narratives seems like a natural choice, of course. Yet, saying that the literary samples are extremely well-chosen almost seems like an understatement because they are so clearly *formative* to Polvinen’s arguments.

This formative role of fictional works is so essential to the practice of 4E cognitive narrative studies it may be worth explaining in a bit more detail. What I mean is that the vast majority of narrative theory leaves the impression that the author has begun with a theory, a system, or perhaps a reaction to previous theories and then cherry-picked the most appropriate narratives to demonstrate their ideas. *Self-Reflective Fiction and 4E Cognition*, by contrast, reads like its theoretical notions have emerged more organically from the interactions between Polvinen’s theoretical expertise and the innovative aspects of the narratives discussed. They are hardly just “target texts.” Rather, Polvinen’s analyses seem to prove both Brian McHale’s claim that speculative fiction can function as literalized narratology (11) and the 4E theories’ conception of cognition as something that arises from the interactions between the cognizer and their environment. If these two notions gain popularity and traction, they might well revolutionize the entire field of literary studies, first, by highlighting SFF as a genre of exceptional (meta)cognitive value, and second, by dissolving the hierarchy between artistic works and the theoretical tools we purportedly use to dissect them.

Overall, the 4E paradigm is a paradigm of two-way interactions – or at least it ought to be. Just as literary theory should be shaped reciprocally by both literature and theory, cognitive literary studies would also be more fruitful as a genuine interdisciplinary dialogue than as the humanist monologue it currently is. While literary studies have adopted useful tools and terminology from cognitive studies, the insights that cognitive studies might gain from its intersections with literary studies remain largely unexplored. Thus, one can only hope that *Self-Reflective Fiction and 4E Cognition* finds its way to the more sciencey side of university campuses and does its part in realizing the optimistic vision the book concludes with: cognitive literary studies has real potential “to generate fruitful hermeneutic circulation around the study of the imagination in both the sciences and the humanities” (151).

Of course, in reading *Self-Reflective Fiction and 4E Cognition*, a skilled metacognizer can also see the roads Polvinen has not taken. Although the title of the book purports that it examines self-reflective “fiction,” its viewpoint is exclusively literary. That is to say, there are hardly any references to fiction in other media, even though the unity of form and content – and thus, the

indivisibility of mimesis and artifice – is arguably more salient in drawn images. On the other hand, some visual media require the reader to physically oscillate their attention between different structures: for instance, reading a comic requires adjusting one’s gaze back and forth, zooming into the individual frames and out to the full layout again. The perceptual apparatus of the reader thus resists the notion of “lamination” in the context of multimodal texts, at least to a degree, and the transmedial applicability of *Self-Reflective Fiction and 4E Cognition* remains debatable, at least for the time being.

Furthermore, while Polvinen is exceptionally observant about the metaphorical language used not only in literature, but also in its study, one could always go even deeper and explore the logic behind the metaphors. For instance, understanding abstract and imaginary things in spatial terms is by no means specific to narratology. It is a general, unavoidable offshoot of the fact that cognition is always embodied and embedded in lifeworlds – that is, 4E cognition is inherently spatial, and the primacy of spatial-visual thinking has even been proved by neuroimaging (see, e.g., Amit et al.). It is no wonder, then, that most writers and readers also understand literature as a kind mental space – or, indeed, a cognitive environment. To quote Mary Oliver, “the poem is a temple – or a green field – a place to enter, and in which to feel. Only in a secondary way is it an intellectual thing – an artifact” (12).

Of course, this does not contradict Polvinen’s main claims that literature is both an imaginary place and an artistic artifact all at once and that we can understand literature and gain something cognitively valuable from it due to this very duality, not in spite of it. Likewise, none of my points of criticism dispute that *Self-Reflective Fiction and 4E Cognition* is an erudite work of cognitive literary theory. While the corrections and additions it makes to state-of-the-art narratology may seem subtle on the surface, they seep deep into the very nature of literary and theoretical imagination. Polvinen’s self-reflective inquiries and speculative investigations should thus be considered far and wide in the ranks of narrative and cognitive theorists alike.

*Biography:* Essi Varis, PhD, has studied narratives across media from various cognitive perspectives since 2013. She is currently working on the final stages of her personal post-doctoral research project titled *Metacognitive Magic Mirrors: How Engaging with Texts and Images Enables and Extends Speculative Thought* (Finnish Cultural Foundation 2020–2025) and serving in the jury of Finland’s national comic book award, Sarjakuva-Finlandia. She also served as an editor-in-chief of *Fafnir* in 2021–2023.

## Works Cited

Amit, Elinor, Caitlyn Hoeflin, Nada Hamzah, and Evelina Fedorenko. “An Asymmetrical Relationship between Verbal and Visual Thinking: Converging Evidence from Behavior and fMRI.” *NeuroImage*, no. 152, 2017, pp. 619–27.

Oliver, Mary. *Upstream: Selected Essays*. Penguin Books, 2016.