How Stories Make Us Feel: Toward an Embodied Narratology

Hannah Chapelle Wojciehowski and Vittorio Gallese

The Rise of Neurohumanism

For the last twenty-five years, a revolution in the science of mind has been underway. Rapid and cascading breakthroughs in the understanding of minds (those of humans and other primates, as well as other animals), the relation of mind to body, and of the perplexing phenomena of human consciousness, memory, empathy, and intersubjectivity have profoundly altered our collective understanding of the nature of our psychic and physical lives, as well as the seamless interface between those two aspects of experience. This revolution has unfolded within a number of academic disciplines, including cognitive neuroscience, psychology, computer science and A.I., linguistics, and philosophy. It has also migrated into literary studies, art history, history, anthropology and other disciplines of the humanities and social sciences. Cross-disciplinary conversations and collaborations have not only enhanced the progress of knowledge-building and understanding, but also encouraged speculative thinking – particularly about what it means to be human (as similar to or distinct from other living beings). Many of these exchanges demonstrate a contagious excitement and enthusiasm that have ignited interest well beyond the walls of the academy, since the questions being broached are of interest to large numbers of people who are eager for information about this new paradigm.

A popular scientific literature addresses that shared curiosity and anxiety, presenting complex information in reader-friendly terms. Myriad books and articles take up questions of human identity, appealing to a pre-existing humanist discourse in order to contextualize recent scientific discoveries. From the other side, a growing number of scholars in the humanities are currently engaging in deeper dialogue and/ or collaborative research with members of the scientific community in order to explore points of conjunction between what were until recently regarded as very separate fields; these humanists have contributed to the rise of cognitive poetics, cognitive historicism, cognitive film studies, and cognitive cultural studies, to name some of the closely related subfields of humanistic research that draw on neuroscience in order to develop new models for the study of texts, whether verbal or visual; of authorship, creativity and intentionality; of reader- or viewer-response; of affect, emotion, and embodiment; of the comparative study of cultures; and of individual or collective identities.

1 A casual survey of some recent titles popularizing the neuroscientific revolution reflects this preoccupation with defining the human: for example, V. S. Ramachandran’s The Tell-Tale Brain: A Neuroscientist’s Quest for What Makes Us Human (2011); Oliver Sacks’ The Mind’s Eye (2010); Marco Iacoboni’s Mirroring People: The New Science of How We Connect With Others (2009); Maryann Wolf’s Proust and the Squid: The Story and Science of the Reading Brain (2008); Giacomo Rizzolatti and Corrado Sinigaglia’s Mirrors in the Brain (2007); and Steven Pinker’s The Stuff of Thought: Language as a Window into Human Nature (2007); as well as Antonio Damasio’s Self Comes to Mind (2010), The Feeling of What Happens (1999), and his now classic Descartes’ Error: Emotion, Reason and the Human Brain (1994).
We will discuss this emerging trend, which we call neurohumanism: namely, the framing of scientific discoveries regarding mind in terms of what it means to be a human, as well as the complex dialogues and interchanges about the nature of the human between members of the global scientific community on the one hand and members of the humanities on the other, with social scientists often poised between the two groups. We shall consider the potential, as well as the risks of neurohumanism, particularly in its most embodied and least “cognitive” forms, for explaining long-standing puzzles of human identity, agency, and sociality, particularly across the science-humanities divide.

In the first half of the essay, we shall trace the rise of antihumanism in the twentieth century, its modulations into posthumanism, and finally the marked return to humanism in the new millennium. We shall also discuss the ways in which the sciences – specifically neuroscience – have contributed to this return, specifically in the theoretical field of cognitive literary studies. In the second half of the essay, we offer our own intervention by proposing a literary narratology based on what we call Feeling of Body (FoB). More specifically, we shall attempt to connect neuroscience and twentieth- and twenty-first century literary studies through an embodied narratology hinging on a functional mechanism – liberated embodied simulation – enabling the FoB and characterizing one important level of our relationships with narrative – namely, our empathic co-feeling with others activated by writings and registered within our own bodies.

From Antihumanism to Cognitive Literary Studies

Until quite recently, appeals to human nature were more likely to be met with skepticism, caution, or even cynicism on the part of many literary theorists, rather than with enthusiasm, fascination, or assent, though this trend is clearly reversing in some quarters. In the last decades of the twentieth century, the humanities, and literary studies in particular, were the site of a pronounced antihumanist discourse – i.e., a sustained critique of notions and ideals concerning human identity that were often associated with the Enlightenment and the so-called ‘project of modernity.’

Antihumanism was one of the overarching themes of post-1968 French philosophy, including that of Jacques Lacan, Roland Barthes, Louis Althusser, Jacques Derrida, and Michel Foucault. Its roots and antecedents could be traced even further back to the philosophies of Marx, Nietzsche, Freud, Weber, Husserl, and Heidegger – each of whom had challenged and critiqued prevailing ideologies of the human, and of political, social, and psychic life as they had been previously understood. As the contemporary political theorist Diana Coole explains, stressing the influence of phenomenologists Husserl and Heidegger, antihumanist thinkers maintained that humanism is itself implicated in an aggressive subjectivist culture that reproduces the hubris and existential impoverishment of the modern age, in which normativity as such succumbs to positivism and nihilism, and Western norms translate into imperialism and colonialism (2007, 28).

Coole makes it clear that antihumanist thought, so prominent in the late twentieth century, and associated most strongly with French neo-Marxism, poststructuralism, and Foucaultian genealogical theory, had its roots in earlier nineteenth- and twentieth-century
critiques. Those earlier challenges were, however, given new force by the May uprisings in 1968, both in Paris and in other parts of the world, and by events leading up to that watershed moment, as Luc Ferry and Alain Renaut have persuasively argued (1990, xvii ff.). Post-'68 Marxists (including Althusser and Pierre Bourdieu) aligned the traditional discourse of humanism with bourgeois ideology, positivism, and the failures and catastrophes of modernity. Meanwhile poststructuralists such as Derrida and Foucault, who constructed their philosophies upon a Nietzschean and Heideggerian framework, folded Marx's view of man's potential mastery of nature and his political environment into their far-reaching critiques of humanist thought. Within all of these discourses, the very word 'humanist' soon became a shorthand term connoting all things retrograde, totalizing, and/or totalitarian—the very essence of false consciousness.

By the 1970s and 80s, antihumanist theories had moved far beyond their original German and French contexts, powerfully influencing academic discourses of the humanities and social sciences throughout Europe, the United States, and elsewhere. By the late 80s, the critique of humanist subjectivity was in full swing.

Liberationist movements such as postcolonial and ethnic studies, feminisms, and the more recent gender and queer studies had, to a greater or lesser extent, incorporated various critiques of subjectivity into their own arguments, and, as a consequence, absorbed the rhetoric of antihumanism. Initially some offshoots of these resistance movements (for example, American feminisms and African-American Studies of the 60s, 70s and early 80s) hewed towards older notions of subjectivity and agency as part of an effort to gain political rights that had previously been withheld or denied. Other groups that were strongly informed by Lacanian psychoanalysis, by Derridean deconstruction, and by discourses of poststructuralism or post-Marxism questioned the strategic value of claiming liberal humanist subjectivity while, in effect, buying into the illusions and ideologies of modernity that several generations of thinkers had so painstakingly challenged. To imagine for oneself or one’s group the possession of that humanist subjectivity was to re-instantiate the hegemonic power relations and oppressive ideologies that each of these critical movements was seeking to dismantle. Literary theorist and deconstructionist Gayatri Spivak, articulating the double-bind of that situation for feminists, theorists of race, and postcolonial critics, suggested a paradoxical solution: namely, taking “the risk of essence” (1989, 129). One could assume a provisional essentialism in order to foster the interests of one’s group, while simultaneously acknowledging within one’s group and among allies the contingency of identity, whether individual or collective. Other critics, finding such suggestions to be

---

2 Ferry and Renaut further argue: “[It] is important to understand that this critique of modern rationality was absolutely inseparable from a critique of the subject (of man) defined as conscience and as will, that is, as man as the author of his acts and ideas. In order to understand this, one must refer back to the considerable trauma represented by the Second World War for European intellectuals. Immediately after the war, in fact, it is no exaggeration to say that ‘civilized societies,’ that is the entire Western world, could legitimately be accused of having engendered, or at least of having been unable to stop, two of the greatest political catastrophes of this century: colonialist imperialism and Nazism” (xii-xiii).

3 Spivak’s proposal, initially offered at a 1982 Modern Language Association meeting in Los Angeles, would become a touchstone for many theorists at that time.
maddeningly hair-splitting, sought to refocus the collective discourse on issues of social justice, without overcomplicating questions of identity, agency, power, and history.4

While these debates on subjectivity flourished throughout the 1980s and into the 1990s, other theorists developed a line of critique that questioned categorical distinctions between human and animal, as well as human and machine – two sides of the same coin. Rejecting the celebratory ‘womanism’ and spirituality of second-wave American feminism, primatologist-turned-literary theorist Donna Haraway famously (and humorously) stated in her 1985 “A Cyborg Manifesto,” “I would rather be a Cyborg than a goddess” (1991, 181). Haraway’s embracing of the non-human, the artificial, and the inorganic, rather than notions of essential femininity, heralded not only a new phase of feminism, but also posthumanism.

In the late 90s, the cultural theorist and erstwhile chemist N. Katherine Hayles further developed this line of posthumanist critique by analyzing the technological supplementation and reconfiguration of human bodies, as well as their various prostheses.5 In her analysis of cybernetics and its impact on our collective understanding of human subjectivity, Hayles wrote, “When the body is revealed as a construct, subject to radical change and redefinition, bodies of knowledge are similarly apt to be seen as constructs, no more inevitable than the organic form that images them” (1999, 85). Hayles highlighted the ways in which the core concept of a “coherent, rational self,” coupled with “the right of that self to autonomy and freedom, and a sense of agency linked with enlightened self-interest” – a foundational concept within liberal humanism – would be challenged by cybernetics and later by informatics (85-86). These discourses complicate traditional divisions upon which liberal subjectivity has been based – divisions such as human/machine, human/animal, and animate/inanimate that also rely on hierarchical distinctions organizing these and other binaries.6

As this brief discussion of antihumanisms (for there are many) and their more recent modulations into posthumanism suggests, critiques of the human have taken many forms over the twentieth century. Since these critiques have occurred within specific disciplinary contexts and among or within specific communities, it is difficult to capture the essence, so to speak, of antihumanism, which remains a protean phenomenon. These

---

4 In her 1987 essay “The Race for Theory, the African-American feminist and Black Liberation activist Barbara Christian challenged the obfuscatory nature of poststructuralist theory, and especially French feminism, while coming back to a clear-cut politics of identity: “[W]e can say that the terms ‘minority’ and ‘discourse’ are located firmly in a Western dualistic or ‘binary’ frame which sees the rest of the world as minor, and tries to convince the rest of the world that it is major, usually through force and then through language, even as it claims many of the ideas that we, its ‘historical’ other, have known and spoken about for so long. For many of us have never conceived of ourselves only as somebody’s other” (2010, 1860).

5 Hayles’ concept of “posthuman” in this work arises out of the idea that through our tools and technologies we have become something more than (or perhaps beyond) human, but also, perhaps, that our bodies have always had supplements of one kind or another. We may note a similar critical turn represented by the Posthumanities series published since 2007 by the University of Minnesota Press (formerly one of the preeminent publishers of poststructuralist theory) and edited by Cary Wolfe.

6 Lest the loss of that concept arouse despair or some related sense of futility, Hayles presents a more optimistic view of posthumanity: “Human being is first of all embodied being, and the complexities of this embodiment mean that human awareness unfolds in ways very different from those of intelligence embodied in cybernetic machines” (283-284). Giving up the older liberal notion of humanity presented for Hayles “the exhilarating prospect of getting out of some of the old boxes and opening up new ways of thinking about what being human means” (285). Hayles’ theorizing of the posthuman body and of distributed cognition is an important contribution to cognitive cultural studies.
and other attempts to slough off liberal humanism, usually in favor of a more examined and empowering, often hybridized, deconstructed, or otherwise paradoxical sense of being in the world, have resulted in some extraordinarily creative accounts of modernity and of history. They have, in sum, been productive in both predictable and surprising ways.

Yet the tide now seems to be turning. In the 2000s, posthumanism has continued to evolve as a theoretical model, though there has been a gradual resurgence of humanism, as well. Critiques of essentialisms, universalisms, totalizations, and nature-over-nurture arguments have, in the space of a generation and a half, begun to cede ground to forms of analysis that once again sound vaguely or overtly humanistic. It is probably not an accident that most of the influential neurohumanist texts published in the last fifteen years, some of the more recent of which were mentioned at the beginning of this essay, were written by scientists or by social scientists highly conversant with neuroscience.

How might we account for this chiasmic phenomenon? Should we assume that it is more acceptable to talk about humans, human nature, and humanism if one is not in the humanities or, ideally, if one is in the sciences, where post-Sixties antihumanism seems to have made the fewest inroads? Do neurohumanist thinkers such Oliver Sacks, Antonio Damasio, and V. S. Ramachandran, for example, speak from a substantially different cultural matrix (or set of matrices) than do humanist scholars? Do they invoke traditional humanist themes in their writings out of a lack of familiarity with the antihumanist tradition described above, or do they merely ignore it in order to talk about questions that most people – especially the reading public – find powerfully interesting and important? In what forms was the humanist tradition presented to these and other neuroscientific thinkers, and how did it shape their research and writing practices? How, in turn, are their ideas contributing to new modes of reading and interpreting within contemporary literary studies – and vice versa?

Clearly humanism is renewing itself and is doing so in precisely the discursive arenas where such questions can be raised without remonstrance – allegations of retrograde or non-progressive politics, for example – namely, outside the humanities as they are often practiced within the academy. It is fascinating to observe that the sciences, and especially certain branches of neuroscience, primatology, ethology, and medicine, have become the repositories of traditional, ‘high’ humanist concepts such as empathy, free will, and creativity, and posit a shared or universal human nature.

In the humanities, meanwhile, the term humanism, together with its attendant themes, is often broached with a certain caution or reluctance, or at least with some presumed need to justify why one might want to ‘go there.’ In a then-unusual 1997 essay entitled “Literary Universals,” Patrick Colm Hogan described the state of affairs as he saw it at the time: “Today,” he wrote,

there is little enthusiasm among humanists for the study of universals. Indeed, it is barely even a concept within the humanities, where the focus of both theory and practice tends to be on ‘difference,’ ‘cultural and historical specificity,’ and so on…. When universalism is mentioned at all in humanistic writing, it is most often denounced as a tool of oppression. (2010a, 37)
The situation Hogan described more a decade ago has changed somewhat in the intervening years, due in large part to the explosion of cognitive poetics and cognitive cultural studies in the 2000s. Neurohumanism is not, however, a label that practitioners of these new theoretical modes in literature, art history, film studies, history and other of the humanities would be likely to embrace. The term humanism still possesses negative connotations for many literary scholars trained during the heyday of poststructuralism and deconstruction, or during the culture wars that ensued during the late 80s and 90s. Moreover, attempts to bring science to bear on literary experience, and on the experience of art in general, are often seen by many in the humanities and fine arts as both reductive and limiting.

Hogan’s cautious appeal to human universals takes the “risk of essence” in a non-deconstructive direction. His argument also raises an important question about what exactly had been set aside – or even repressed – during the peak years of deconstruction and since. Arguably humanism served as a straw man in many critiques, representing at times patriarchy, white hegemony, colonialism, racism, capitalism, bourgeois ideology, and homophobia, as well as many other modes of exploitation or oppression. Simultaneously, humanism performed a quite different function, for it could also be linked to emotionality, naiveté, irrationality, and, for want of a better word, the squishiness of an earlier worldview that seemed silly or vulnerable to attack – in sum, not strong enough to defend humanity against itself. The ultimate shifting signifier, humanism may thus have represented a worldview that was discarded as “effeminate” (though that connection seems not to have been made consciously in most cases), and that was increasingly derided in the very era that women, as well as other once excluded (hence arguably feminized) groups, began to enter the academy in larger numbers than ever before.

The new fields of cognitive poetics, cognitive cultural studies and cognitive historicism, which began to emerge in the 1990s and which have grown exponentially in the 2000s and 2010s, may signal not so much a return to humanism as a set of critical perspectives deriving from disciplines that were less strongly influenced by poststructuralism – especially structural linguistics and narratology – and from cognitive psychology and neuroscience. Leaders in this new set of fields include Patrick Colm Hogan, already mentioned, and Mary Thomas Crane, Suzanne Keen, Alan Richardson, Ellen Spolsky, Reuven Tsur, Peter Stockwell and Lisa Zunshine. Taken collectively, their works offer new purchase on a wide range of questions and topics also encompassed by traditional humanism – e.g., emotions, affects, and empathy; the relation of mind and body (usually considered monistically), reason and action; aesthetics and ethics. Rhetoric, poetics, narrative structure and emplotment, authorial intention and reader response – these and many other staple topics of literary studies throughout the ages have found new

7 We coin the term “neurohumanism” with a certain irony, precisely because of the pronounced history of antihumanism that has dominated scholarship in these fields since the Sixties. However, to others outside of the humanities, neurohumanism may have a different and more positive valence.

8 Representative works by these authors include the following: Hogan, The Mind and its Stories (2003); Crane, Shakespeare’s Brain (2001); Keen, Empathy and the Novel (2007); Richardson, The Neural Sublime (2010); Spolsky, Word vs. Image (2007); Stockwell, Cognitive Poetics (2002); Tsur, Toward a Theory of Cognitive Poetics (1992/2008); and Zunshine, Why We Read Fiction (2006).
expression in the burgeoning field(s) of cognitive literary studies, which constitute a counterpart to scientific neurohumanism, though the modifier “cognitive” signals the reach of these new discourses in the opposite direction across the humanities/science divide — that is, toward cognitive neuroscience and psychology. Frequently, though not invariably, cognitive literary theory draws on so-called Theory of Mind (ToM), the notion that we can reconstruct the minds and intentions of other people through our own mental meta-representational processes. In one noteworthy analysis of a passage from Virginia Woolf’s *Mrs. Dalloway*, Lisa Zunshine provides an account of Theory of Mind at work in the novel, by tracing embedded intentionalities up to a dizzying sixth level (2006, 31-36). Mind-reading, Zunshine and others have argued, is what makes fiction profoundly interesting to readers, for it hones our meta-representational cognitive abilities.

Yet such accounts of mind-reading, compelling as they may be, run the risk of reintroducing a backdoor Cartesianism to our understanding of literary texts, a separation of mind and body into distinct divisions and an implicit privileging of one over the other. What we shall ultimately propose, then, as a complement to Theory of Mind, is the Feeling of Body (FoB), its possible links to the experience of narrative, and its relation to cognitive, as well as affective literary theory (in some cases a subset of recent cognitive literary theory). We will argue that FoB is the outcome of a basic functional mechanism instantiated by our brain-body system, Embodied Simulation, enabling a more direct and less cognitively-mediated access to the world of others. Embodied simulation mediates the capacity to share the meaning of actions, basic motor intentions, feelings, and emotions with others, thus grounding our identification with and connectedness to others. According to this hypothesis, intersubjectivity should be viewed first and foremost as intercorporeity. Thus, FoB is to be conceived as a crucial ingredient of our relationship with fictional narratives.

Partly in response to the highly “cognitive” approach to reading texts noted above, other critics have turned to the study of emotions in literature, whether of the affective responses of readers, affects encoded within texts, and/or the affects within the public sphere. As Hogan writes, “we have tended to treat narrators as Vulcan-like. But a pure Vulcan narrator would not narrate. Put in the most basic way — if it talks, it feels. Narrators necessarily have emotions” (2010b, 77). In the second part of the paper we will make explicit not only why we believe Hogan is right, but will also broaden the crucial

---

9 On different expressions of the “affective turn” in literary studies, see e.g., Massumi (2002); Sedgwick, (2003); Hogan (2011); and Staiger, Cvetkovich and Reynolds (2010). In “The Turn to Affect: A Critique” (2011), Ruth Leys cautions against separating affect and cognition, and of privileging the former, both of which she takes as the premier features of affective criticism. We do not see one as exclusive of the other, but rather each as inextricably related to the other.

10 Mr. Spock, the famous character of the Sci-Fi saga Star trek, comes from Vulcan, and although looking like a human (but for his sharp ears), he is a purely logical creature. Interestingly enough, Gallese wrote: “in contrast to Mr. Spock, the famous character of the Star Trek saga, our social mental skills are not confined to a declarative third-person perspective. We are not alienated from the actions, emotions and sensations of others, because we entertain a much richer and affectively nuanced perspective of what other individuals do, experience, and feel. What makes this possible is the fact that *we own* those same actions, emotions, and sensations” (2005a, 31).
role of the body for narrative theory, by including in the notion of FoB a much wider domain, encompassing actions and sensations.\textsuperscript{11}

First, however, we shall explore the term “cognitive” within neuroscience, psychology, and philosophy of mind, where that term conveys a somewhat different set of meanings.

*From Classic Cognitivism to Embodied Cognition*

Scientific and philosophical debates over the “cognitive” – as over consciousness, mind, and simulation – have been a significant influence on the new branches of literary studies mentioned above; hence, it is crucial to establish the nature of those debates in their original contexts, and also to consider what is and is not osmosing across the membranes dividing the sciences and the humanities.

As we stated in our introduction, the second half of the twentieth century was marked by the flourishing progress of cognitive neuroscience, made possible through the newly developed brain imaging technologies, like the fMRI,\textsuperscript{12} enabling the non-invasive study of the human brain. Neuroscience started investigating domains such as intersubjectivity, the self, empathy, decision-making, ethics, aesthetics, economy, etc. As Giovanni Frazzetto and Suzanne Anker recently wrote, various domains of knowledge have acquired a “neuro” dimension. We are witnessing “the rise of a neuroculture (or neurocultures), in which neuroknowledge partakes in our daily lives, social practices, and intellectual discourses…. As part of this transformation, ideas, images and concepts of neuroscience are increasingly assimilated into the cultural imaginary” (2007, 815). Furthermore, the bidirectional interaction between the brain and cultural experience is today the object of investigation of a new emerging field, cultural neuroscience (Chiao and Harada 2008; Losin et al. 2010).

The media hype on supposedly new disciplines like Neuroaesthetics, Neuroethics, Neuroeconomics and the like has stirred irritated reactions, especially within the camp of cognitive science. A recent book entitled *Neuro-mania: il cervello non spiega chi siamo* (2009), by the Italian cognitive scientists Paolo Legrenzi and Carlo Umiltà, strongly criticizes the neurological positivism and determinism generated by the acritical application of the brain-imaging approach to practically all dimensions defining the human condition. These authors are right in emphasizing the limits intrinsic to the brain-imaging approach, which does not directly measure neural activity, but can only estimate it indirectly by detecting how blood flow varies across different brain regions during perceptual, motor, or cognitive tasks. Such criticisms, however, miss the point by

\textsuperscript{11} In doing so, we are elaborating on the models for reception theory and reader-response criticism developed by Wolfgang Iser, Hans Robert Jauss, Stanley Fish and many others. See, e.g., Iser (1974) and (1978); Jauss (1982a and 1982b); Fish (1980); and Tompkins (1981). We would add a component of embodiment to these theories by highlighting the neurological bases of readers’ experiences that sublend cultural and historical responses to texts.

\textsuperscript{12} Functional Magnetic Resonance Imaging, or fMRI, is a technique of mapping regional changes of blood flow within the brain, relating it to neural brain activity. This approach, which was developed in the late 1980s, has led to a greater understanding of the functional properties of brain regions involved in a variety of cognitive, motor and perceptual tasks, as well as in many other aspects of human cognition.
advocating for neuroscience the limited and limiting role of localizing in the brain the psychological mechanisms identified by cognitive science. Nevertheless, cognitive neuroscience can challenge many of the psychological mechanisms taken for granted by classic cognitive science. The crucial role of action in perception, as epitomized by the discovery of mirror neurons (see below), is one good example of how neuroscience can radically change old views.

Neuroscientific investigation has recently challenged many assumptions of classic cognitive science on the human mind. The truth is that there is no need to envisage new disciplines with the prefix neuro-. A much more important issue is whether cognitive neuroscience can shed new light on the most distinctive aspects defining the human condition, like art, creativity, and aesthetics. In our opinion the answer is positive.

Before we attempt to answer those questions, however, we must emphasize that cognitive neuroscience is a methodological approach whose results are deeply shaped by the theoretical assumptions framing it. Vast areas of cognitive neuroscience are still deeply influenced by classic cognitivism, on the one hand, and by evolutionary psychology, on the other. Classic cognitive science heralds a solipsistic account of the mind. By solipsism we mean that focusing on the single individual’s mind is all that is required to define what a mind is and how it works. The picture of the mind conveyed by classic cognitive science is that of a functional system whose processes can be described in terms of manipulations of informational symbols according to a set of formal syntactic rules. The late philosopher Susan Hurley characterized this approach with the metaphor of the “Classical Sandwich.” In her book Consciousness in Action Hurley wrote:

A view of perception and action as separate input and output systems complements a view of thought and cognition as “central” and in turn separate from the “peripheral” input and output systems. The virtual processing of cognition is seen as central, even if its implementation is distributed; input to it is provided by perception, and it issues output that generates action. The subpersonal underpinnings of the mind are conceived as vertically modular, with cognition interfacing between perception and action. (1998, 20)

In other words, classic cognitivism conceives of action perception and cognition as separate modular domains, neglecting both the intertwined character of perception and action and their crucial contribution to cognitive processes. According to classic cognitivism, representations are symbolic in nature, and thinking can be reduced to computation. Consequently, the understanding of other minds is conceived solely as a predicative, inferential, theory-like process.

Evolutionary psychology’s defining metaphor, in contrast to the computer metaphor underlying classic cognitivism, is that of the Swiss-Army knife. According to this

---

perspective, the human mind is a collection of cognitive modules, each of which has been selected in the course of evolution because of its specific adaptive values. According to Leda Cosmides and John Tooby, “All normal human minds reliably develop a standard collection of reasoning and regulatory circuits that are functionally specialized and, frequently, are designed to operate specifically within a particular domain (for example, sexual behavior, foods, navigation). That is, they are often domain specific” (1997, 75). The five principles they list as capable of explaining any topic in psychology are Principle 1: The brain is a physical system functioning as a computer and has circuits designed to generate behavior appropriate to environmental circumstances; Principle 2: Neural circuits were designed by natural selection to solve problems that ancestors faced during the species’ evolutionary history; Principle 3: Most of what goes on in the mind is hidden; thus, most problems that seem easy to solve are actually very difficult to solve - they require very complicated circuitry; Principle 4: Different neural circuits are specialized for solving different adaptive problems; and Principle 5: Modern skulls house Stone Age minds (75-84). A linguistic module, a Theory of Mind module, and the like would be the cognitive tools by means of which individual human beings express a variety of instincts, like that of language (Pinker 1994).

By explicitly or implicitly conforming to these views when investigating social cognition, cognitive neuroscience during the last twenty years has been mainly focused on finding a house in the human brain for the above-mentioned cognitive modules. This standard approach leaves itself open to what Bennett and Hacker have called the “mereological fallacy” – that is, attributing to one part of a living organism – e.g. the functional properties of the nervous system – characteristics that are proper to the entire organism as a whole (2003, 68). The point is that the ability to understand mental states – and intersubjectivity – the sharing of subjective states by two or more individuals – are competences uniquely describable at the personal level, and therefore are not entirely reducible to the sub-personal activation of neural networks in the brain, hypothetically specialized in mind-reading, as too many neuroscientists nowadays think (and, following their lead, some cognitive narratologists). Indeed neurons are not epistemic agents. The only things neurons “know” about the world are the ions constantly flowing through their membranes. In contrast, mentalization and intersubjectivity are personal-level properties of individuals (Gallese 2009, 165).

This widespread epistemological attitude, which may be called cognitivist ontologic reductionism, is often combined with a reliance on brain imaging techniques like fMRI as the sole method of investigation. The problem is that fMRI studies, if not supported by a detailed phenomenological analysis of the investigated perceptive, motor, and cognitive processes, and – even more importantly – if not interpreted on the basis of the direct study of the activity of single neurons in the animal model, lose much of their heuristic power. The heuristic power of this approach is further reduced by the instrumental use

---

15 Yet, as Louis Menand wrote in his devastating review of Pinker’s The Blank Slate, “To say that music is the product of a gene for ‘art-making,’ naturally selected to impress potential mates – which is one of the things Pinker believes – is to say absolutely nothing about what makes any particular piece of music significant to human beings. No doubt Wagner wished to impress potential mates; who does not? It is a long way from there to ‘Parsifal'” (2002, 97).
16 By means of fMRIs, one can establish a correlation between a given task and the activation of a given brain area. However, the neurophysiological mechanisms explaining why activity in a given brain area
of empirical data to validate a preconceived model of the mind, considered truly \textit{a priori}. Such a model, most of the time, is the one proposed by classic cognitive science, according to which mental states are theoretical states of a commonsense psychological theory called Folk Psychology.\footnote{Also known as vernacular or common-sense psychology, Folk Psychology may be considered the set of beliefs or assumptions that shape the language we use to discuss human psychology.} According to this view, understanding others squares with the possibility of attributing to others propositional attitudes like beliefs, desires and intentions. It is highly questionable that this model fully captures the expression and functional architecture of the human mind. As we shall argue, these debates within the global community of neuroscientists, which may be more or less “cognitive” in their emphases, and also more and less holistic in their assumptions about human cognition, do not always get transmitted to literary scholars drawing on recent neurocognitive theory. What comes across, in other words, is a model of mind more “cognitive” and/or disembodied than it needs to be.

The truth is that neuroscientists do not yet have a clear model of how humans understand each other. At present, we have a series of brain imaging studies showing the activation of a set of cortical brain regions during explicit mentalizing tasks, like attributing false beliefs to others. No one to date has been able to provide a convincing explanation about why these specific brain regions do activate during mentalization, besides the tautological statement that mind-reading (which we will discuss further below) is implemented there. In fact, the mind-reading specificity of the activation of these brain regions is highly debatable. fMRI studies present yet another text to be interpreted; the information they convey is at times highly ambiguous and/or subject to multiple interpretations.

Classic cognitivism, like a dead star, still emits light. Much of what psychology students learn about how the mind supposedly works is still informed by this paradigm. Furthermore, many neuroscientists unquestioningly accept this model of the mind and confine or limit their empirical investigation to a mere localization in the brain of processes and mechanisms they accept \textit{a priori}. We believe, though, that we are in the middle of another paradigm shift. A new scientific approach to the study of the human condition is gaining momentum: it is the so-called “embodied cognition” approach.\footnote{See Lakoff and Johnson (1980 and 1999); Lakoff (1987); Clark (2010); Barsalou (1998); Anderson (2003); Gallese (2000, 2001, 2003); Barrett and Henzi (2004); and Niedenthal et al. (2005).} As George Lakoff already envisaged many years ago: “Thought is embodied, that is, the structures used to put together our conceptual systems grow out of bodily experience and make sense in terms of it; moreover, the core of our conceptual systems is directly grounded in perception, body movements, and experience of a physical and social character” (1987, xiv). Lakoff’s embodied conception of thought and of human conceptual systems anticipated discoveries in primate neuroscience just a few years later.

The discovery of mirror neurons in the macaque monkey brain during the early 1990s (Gallese et al. 1996; Rizzolatti et al. 1996; see also Gallese and Goldman 1998) and the subsequent discovery of mirroring mechanisms in the human brain (see Gallese...
2001, 2003, 2007; Gallese et al. 2004; Rizzolatti and Sinigaglia 2007) suggested that a more direct access to the meaning of others’ behavior is available – more direct, that is, than the explicit attribution of the propositional attitudes of Folk Psychology characterizing standard accounts of mind-reading, as we shall explain. The discovery of mirror neurons provides a new empirically based notion of intersubjectivity, viewed first and foremost as *intercorporeity* – the mutual resonance of intentionally meaningful sensory-motor behaviors – as the main source of knowledge we directly gather about others (Gallese 2007, 2009). Such discovery highlights the role of the body in social cognition.

*Mirroring Mechanisms and Embodied Simulation*

What exactly are mirror neurons? Mirror neurons are premotor neurons first described in the most anterior sector of macaque monkey ventral premotor cortex, area F5 (di Pellegrino et al. 1992; Gallese et al. 1996; Rizzolatti et al. 1996). They typically discharge both when a motor act is executed and when it is observed being performed by someone else. The same motor neuron that is activated when the monkey grasps a peanut is also activated when the monkey observes another individual performing the same motor act. Neurons with similar properties were also discovered in a sector of the inferior parietal lobule reciprocally connected with premotor area F5 (Gallese et al. 2002; Fogassi et al. 2005; Rozzi et al. 2008). Thus, the functional properties of mirror neurons are not the exclusive prerogative of a single location in the brain, but characterize a parieto-premotor cortical network.

Observing an action causes in the observer the automatic activation of the same neural mechanism that is triggered by executing that action oneself. For the first time ever, a neural mechanism allowing a direct mapping between the visual description of a motor act and its execution has been identified. The Parma group proposed that this neural mapping mechanism provides an economical solution to the problem of translating the results of the visual analysis of an observed movement – in principle, devoid of meaning for the observer – into something that the observer is directly able to understand because it can be directly mapped onto the observer’s motor representation, that is, on the same neural network the observer normally activates when executing the same motor act (Gallese et al. 1996; Rizzolatti et al. 1996).

It has been proposed that mirror neurons allow a *direct* form of action understanding (Gallese et al. 1996; Rizzolatti et al. 1996) through a mechanism of embodied simulation (Gallese 2005a, 2005b, 2006; Gallese et al. 2009). Such a form of action understanding is considered direct to the extent that it does not require any inference by analogy or other more cognitively sophisticated and explicit forms of mentalization. When we see someone acting or expressing a given emotional or somatosensory state, we can directly grasp its content without the need to reason explicitly about it.

---

19 Embodied simulation is conceived of as a basic functional mechanism of our brain, enabling not only a direct bodily access to the actions, emotions and sensations of others, but also the possibility to imagine similar self- and other-related contents.
Two decades of research have demonstrated the existence within the human brain of a mechanism that directly maps action perception and execution, defined as the Mirror Mechanism (MM) (for review, see Rizzolatti et al. 2001; Gallese 2003, 2006; Gallese et al. 2004; Rizzolatti and Sinigaglia 2007). During action observation, there is activation of the premotor and posterior parietal areas, the likely human homologues of the areas of macaque brains in which mirror neurons were originally described. The mirroring mechanism for actions in humans is somatotopically organized; in other words, the same cortical regions that are normally active when we execute mouth-, hand-, and foot-related acts are also activated when we observe the same motor acts executed by others. (We shall return to this subject in our literary analysis.) Watching someone grasping a cup of coffee, biting an apple, or kicking a football activates the same cortical regions of our brain that would be activated if we were doing the same thing. The Mirror Mechanism instantiated by the parieto-premotor areas can also map basic motor intentions, like eating, drinking, and putting objects away (Iacoboni et al. 2005; Cattaneo et al. 2008). This high level of motor abstraction generates the possibility of executing, hence also of recognizing in the perceptual domain, an orderly sequence of motor acts appropriately concatenated to accomplish a distal goal, as when we reach for and grasp a cup of coffee in order to drink from it. This evidence shows that notions like motor goals and motor intentions\(^{20}\) are integral part of the functional architecture of primates’ cortical motor systems.

Most important for the purpose of our paper, the premotor cortex exhibiting the Mirror Mechanism is also involved in processing action-related words and sentences, suggesting that mirror neurons, together with other parts of the sensory-motor system, could play a relevant role in the semantics of language (see Gallese and Lakoff 2005; Gallese 2007, 2008b; Glenberg and Gallese 2011). It should be added that our understanding of linguistic expressions is not solely an epistemic attitude; it is a way of being. Our way of being, in turn, depends on what we act, how we do it, and how the world responds to us. According to Paul Ricoeur, language is first and foremost discourse, and therefore the “mimetic bond between the act of saying and effective action is never completely severed” (1991, XIV). By means of discourse, language acquires a situated world. It is in discourse that all meanings are transferred, hence “discourse not only has a world but has an other, another person, an interlocutor to whom it is addressed” (78). The hermeneutic development of phenomenology in Ricoeur’s approach connects intentionality to meaning: the logical sense of language must be grounded in the broader notion of meaning that is coextensive with the notion of intentionality (40). This perspective, if complemented with the recent discoveries of neuroscience on the relationship between body and language, corroborate the potentialities of applying the Feeling of Body to narrative theory.

Other mirroring mechanisms seem to be involved with our capacity to share emotions and sensations with others (Gallese 2001, 2003, 2006; de Vignemont and Singer 2006). When we perceive others expressing a given basic emotion such as disgust,

\(^{20}\) By motor goal we imply the purpose of a given motor act, like moving the hand in order to grasp an apple. By motor intention we imply the purpose of a sequence of properly connected motor acts, like grasping an apple to eat it.
some of the same brain areas are activated as when we subjectively experience the same emotion (Wicker et al. 2003).  

Embodied simulation as it is described here is quite different from standard accounts of the Simulation Theory of mind-reading. Embodied simulation is a mandatory, pre-rational, non-introspective process – that is, a physical, and not simply ‘mental’ experience of the mind, emotions, lived experiences and motor intentions of other people. Embodied simulation challenges the notion that interpersonal understanding consists solely of our explicitly attributing to others propositional attitudes like beliefs and desires, which we map as symbolic representations within our own minds. Parallel to the detached third-person sensory description of our social world, embodied simulation creates internal non-linguistic “representations” of the body-states associated with actions, emotions, and sensations within the observer, as if he or she were performing a similar action or experiencing a similar emotion or sensation. This is what the Feeling of Body (FoB) amounts to. By means of the neural format we share with other human beings, and, to an extent, with some animals, as well, we can map others’ actions onto our own motor system, as well as others’ emotions and sensations onto our own viscero-motor and somatosensory systems. It has been proposed that empathy is rooted in embodied simulation (Gallese 2003, 2005a, 2006). Consequently, the FoB is not to be uniquely conceived of as a mere sensing of how our body reacts to external stimuli. It is a bodily way of making sense of our social world.

Our capacity to empathize with others (indeed, a classic topos of humanist thought) is most likely mediated by the embodied simulation mechanisms described above – that is, by the activation of the same neural circuits underpinning our own actions, emotional and sensory experiences (see Gallese 2005a and b, 2006, 2009; Gallese et al. 2004). Following this perspective, empathy may be conceived as the outcome of our natural tendency to experience interpersonal relations first and foremost at the implicit level of intercorporeity. Our interpersonal relations – both in daily life as well as with fictional characters – are marked by our bodily involvement (the FoB) with the actions, emotions and sensations acted and expressed by others.

---

21 This suggests that Mirror Mechanisms could be relevant both to emotional contagion (a notion describing the emotions we seem to catch from others, and which impose themselves on us as unexpected feelings) as well as to our capacity to empathize with others. Empathy, different from emotional contagion, entails the understanding of others’ behaviors and feelings as theirs. In fact, in spite of a common focus of activation in the anterior insula, no matter whose disgust is at stake, different cortical areas activate when disgust is subjectively experienced with respect to when it is only observed in the facial expression of someone else facial expression of someone else (Jabbi et al. 2008). Similar direct matching mechanisms have been described for the perception of pain (Hutchison et al. 1999; Singer et al. 2004; Avenanti et al. 2005; Jackson et al. 2005; Botvinick et al. 2005) and touch (Keysers et al. 2004; Blakemore et al. 2005; Ebisch et al. 2008, 2010).

22 According to Goldman, mind-reading “is to form a judgment, belief, or representation that a designated person occupies or undergoes (in the past, present, or future) a specified mental state or experience” (2009a, 312). Hence, simulation-based mind-reading consists of putting oneself into the mental shoes of others. Simulator and target share the same mind/brain state; such a state is then projected onto the target (2006). According to this view, simulation involves conscious, self-aware thought. See also Goldman, (2009b).
So far we have seen that our brain-body system re-uses part of its neural resources to map others’ behavior. When we witness actions performed by others, we simulate them by activating our own motor system. Similarly, by activating other cortical regions we re-use our affective and somato-sensory neural circuits to map the emotional and somato-sensory experiences of others.23

We should add that the Mirror Mechanism is just one instantiation of embodied simulation, where the simulation process is triggered by a perception (e.g., observing someone doing something, expressing an emotion, or undergoing a somato-sensory stimulation). Embodied simulation, though, can also occur when we imagine doing or perceiving something.24 The power of our imagination is seemingly infinite. Indeed, mental imagery has long been considered one of the most characteristic aspects of the human mind in that it was thought to best epitomize its disembodied nature. However, neuroscience has demonstrated that visual imagining shares with visual perception several features (for a review, see Farah 2000; Kosslyn and Thompson 2000). For example, the time employed to scan a visual scene is matched by the time employed to mentally imagine the same scene (Kosslyn et al. 1978). Furthermore, brain-imaging studies show that when we imagine a visual scene, we activate the same visual regions of our brain normally active when we actually perceive the same visual scene (Farah 1989; Kosslyn et al. 1993; Kosslyn 1994), including the primary visual cortex (LeBihan et al. 1993).

As with visual imagery, motor imagery also shares many features with its actual counterpart. Mentally rehearsing a physical exercise induces an increase of muscle strength to some extent comparable to that attained by a real exercise (Yue and Cole 1992). When we imagine performing a given action, several bodily parameters behave as if we were actually executing the same action (Decety et al. 1989). Jean Decety and collaborators (1991) have shown that heartbeat and breathing frequency increase during motor imagery of physical exercise. As with real physical exercise, these parameters linearly increase with the increase of the imagined effort. Finally, brain-imaging experiments have shown that motor imagery and real action both activate a common network of brain motor centers such as the primary motor cortex, the premotor cortex, the supplementary motor area (SMA), the basal ganglia and the cerebellum (Roland et al. 1993).

23 It should be emphasized that most of the abovementioned empirical results were obtained by showing to volunteers lying in a fMRI scanner videos portraying actions, emotions or sensations executed and experienced by other characters. In a sense, the investigation of how our brain-body system relates to and processes reality is carried out by employing a fictional rendition of real-life situations. The very few experiments that specifically addressed the issue of the relationship between the degree of activation of the mirror neuron mechanism and the real or virtual nature of the agent being observed showed that both situations activate the same parieto-premotor cortical networks. However, actions performed by a real agent in front of the observer elicit stronger responses with respect to videos portraying those same actions.

24 Embodied simulation may be compared to Antonio Damasio’s “as-if body loop,” which enables us to feel an emotional state “as if the body were being activated and modified,” rather than solely from “actual” states. (1995, 154 et passim). In later iterations of this idea Damasio incorporates the Mirror Neuron research of the Parma Group and other researchers (see, e.g., 2003, 115-121, 148-150; and 2010, 102-104, 120-121. In this essay we expand Damasio’s model by exploring the conditions for embodied simulation in response to narrative.
Taken together, these data show that typical human cognitive activities such as visual and motor mental imagery, far from being exclusively symbolic and propositional, rely and depend upon the activation of sensory-motor brain regions. Visual imagery is somehow equivalent to simulating an actual visual experience, and motor imagery is also somehow equivalent to simulating an actual motor experience. Thus, motor and visual imagery do qualify as further forms of embodied simulation, since they imply re-using our motor or visual neural apparatus to imagine things and situations we are not actually doing or perceiving.

It appears therefore that the border between real and fictional worlds is much more blurred than one would expect. This finding opens interesting scenarios for a neurocognitive approach to art, in general, and to narrative, in particular. As the Italian philosopher Alfonso Iacono recently proposed, entering into the fictional world of art implies the inhabiting of an intermediate world whose fictional character is naturalized, henceforth it acquires a natural character, in spite of its artificial nature:

(Translation ours)

How can we use the neuroscientific evidence discussed in the present paper to shed new light on how we process, make sense of, and appreciate the fictional worlds portrayed in novels, poems, theatrical plays, and movies? We propose that embodied simulation can be relevant to our experience of narrative for two reasons: First, because of the Feeling of Body triggered by narrated characters and situations with whom we identify by means of the mirroring mechanisms they evoke. In such a way, embodied simulation generates the
peculiar seeing-as that plays a peculiar role in our aesthetic experience. Second, because of the bodily memories and imaginative associations the narrated material can awake in our readers’ minds, without the need to reflect upon them explicitly.

There is a further aspect characterizing embodied simulation when driven by our immersion into the fictional worlds of art in general, and of narrative in particular, with respect to when this functional mechanism is activated by daily-life situations. In fact, artistic fiction is often more powerful than real life in evoking our emotional engagement and empathic involvement. Why? Perhaps because in aesthetic experience we can temporarily suspend our grip on the world of our daily occupations. We liberate new energies and put them into the service of a new dimension that, paradoxically, can be more vivid than prosaic reality. The aesthetic experience of art works, more than a suspension of disbelief, can be thus interpreted as a sort of “liberated embodied simulation.” When reading a novel, looking at a visual art work, or attending a theatrical play or a movie, our embodied simulation becomes liberated, that is, it is freed from the burden of modeling our actual presence in daily life. We look at art from a safe distance from which our being open to the world is magnified. In a sense, to appreciate art means leaving the world behind in order to grasp it more fully (Gallese 2008b, 2010, 2011). Through an immersive state in which our attention is focused on the narrated virtual world, we can fully deploy our simulative resources, letting our defensive guard against daily reality slip for a while.

Another important element of liberated simulation consists in the fact that when we read a novel (or watch a movie, a theatrical play, or behold a painting), we do it almost completely still. It is a sort of regression to, or, better, a neotenic retention of developmental time when our interactions with the world were almost exclusively mediated by a simulative perception of the events, actions, and emotions populating our social landscape. As recently pointed out by the Italian historian and mass-mediaologist Peppino Ortoleva, the aesthetic experience induced by different forms of art, like theater or cinema, consists of a sort of emotional transfer between actors and spectators that, being forced to inaction, are more open to feelings and emotions (2010). We believe that the same logic applies to reading novels, stories, poems, etc. When we read we not

---

25 Literary theorist Elaine Scarry offers an interesting perspective on this problem in *Dreaming by the Book*, explaining: “I have consistently written here about the way in which a poem or a novel is a set of instructions for mental composition – in something of the way that a musical score provides a set of directions for how to reconstruct the music the composer heard long ago in his or her own head. The “instructional” character is key, because it allows the image to seem to come into being by an agency not one’s own.” (2001, 244).

26 Embodied simulation has been discussed in relation to the aesthetic experience of visual arts (see Freedberg and Gallese 2007; Gallese and Freedberg 2007; Stafford 2007; Di Dio and Gallese 2009; Gallese 2010, 2011; Gallese and Di Dio 2011).

27 Neoteny is the retention of juvenile features in the adult animal. Neoteny in our species likely greatly contributed to the development of psychological traits like the ability to learn, curiosity, playfulness, and sociability.

28 As proposed by Lepage and Théoret (2007), the development of the Mirror Mechanism can be conceptualized as a process whereby the child learns to refrain from acting out the automatic matching mechanism linking action perception and execution. Such development can be viewed as a transition from mandatory re-enactment to a covert simulation of the observed motor acts, most likely through the maturation of prefrontal inhibiting mechanisms. See also Gallese et al. (2009, 106).

29 A similar perspective on stillness was proposed by Denis Diderot in *The Paradox of Acting* (1883) and, more recently by Edgard Morin in *The Cinema, or the Imaginary Man* (2005).
only entirely focus our attention on the literary work, but at the same time our stillness enables us to deploy fully our embodied simulation resources at the service of our immersive relationship with the narrated characters.\textsuperscript{30} This is why the FoB generated by fiction is often more powerful than that triggered by our daily interactions with the world. Our pleasure in novel reading is thus likely also driven by this sense of safe intimacy with a world we not only imagine, but also literally embody.

A similar perspective, as we will argue in the next section, can be applied to the creative process of the author. Liberated embodied simulation hence provides a potentially unified level of description of both author’s and reader’s relation with the text. As the writer Siri Hustvedt recently explained:

Fictions are born of the same faculty that transmutes experience into the narratives we remember explicitly but which are formed unconsciously…. We do not have to be Cartesian dualists to think of imagination as a bridge between a timeless core of sensorimotor affective self and the fully self-conscious, reasoning and/or narrating linguistic cultural self, rooted in the subjective-intersubjective realities of time and space. Writing fiction, creating an imaginary world, is, it seems, rather like remembering what never happened. (2011)

Cognitive Neuroscience, Narrative and Liberated Embodied Simulation

Simulation theory in various forms is beginning to make its way into the study of narratology.\textsuperscript{31} In the remainder of this essay, we shall push simulation further in the direction of embodied cognition by analyzing the FoB within the experience of narrative. What we would like to show at the end of our essay is how a practice of embodied narratology might work – that is, liberated embodied simulation as envisioned by an author, encoded into a text, and experienced by the reader, partly though not exclusively, at the author’s prompting. Each of these three components plays a role within this narratology. We take as our test cases two well-known literary texts in order to suggest the potentials of such a critical framework, though a longer treatment is required to develop these ideas more fully.

The first is Virginia Woolf’s 1925 novel Mrs. Dalloway; the second, Sonnet I of Dante’s 1295 Vita nuova. As unlike as any two works are likely to be, these texts nevertheless provide an excellent, if contrastive, starting point for conceptualizing embodied narratology. We begin with one of the climactic scenes in Woolf’s well-known novel (much beloved by cognitive literary critics) in order to explore her authorial use of

\textsuperscript{30} Admittedly not all novels adhere to the conventions of mimesis or realism, though they may still evoke feelings of body in a variety of ways.

\textsuperscript{31} Adapting Alvin Goldman’s ideas, Blakey Vermeule incorporates simulation theory in her discussion of narratology, hewing closely, however, to Theory of Mind and evolutionary psychology, rather than embodied experience (2010, 39-48). Hogan, following Keith Oatley (1999), also touches upon simulation in order to foreground the pleasure deriving from simulating painful situations in fiction, or vice versa (2011, 28-29; 68-69, et passim). See also Salgaro (2009) and Ceserani (2010).
embodied simulation and to explore its potential effects on readers. Understanding the role of the Mirror Mechanism of both author and reader actually explains this scene in a fundamentally new way. In the scene, which occurs near the end of the novel, the female protagonist Clarissa Dalloway overhears a disturbing conversation that takes place at a party she is holding at her home, in which the suicide of a young, highly traumatized war veteran in London earlier that day is being discussed. The hostess thinks to herself:

What business had the Bradshaws to talk of death at her party? A young man had killed himself. And they talked of it at her party – the Bradshaws, talked of death. He had killed himself – but how? Always her body went through it first, when she was told, suddenly, of an accident; her dress flamed, her body burnt. He had thrown himself from a window. Up had flashed the ground; through him, blundering, bruising, went the rusty spikes. There he lay with a thud, thud, thud in his brain, and then a suffocation of blackness. So she saw it. But why had he done it? And the Bradshaws talked of it at her party! (2005, 179)

While the reader may initially perceive Clarissa Dalloway’s response to the news of Septimus Smith’s death as highly narcissistic or non-empathic, it is essential to note that her first reaction is an immediate and involuntary one. “Always her body went through it first, when she was told, suddenly, of an accident,” Woolf writes, thus presenting a convincing picture of embodied simulation. Interestingly, the rest of the line presents her embodied simulation of another accident: “her dress flamed, her body burnt.” Woolf then switches to a graphic description of Septimus’ leap to his death, not supplied in an earlier description of his jump from the window of his apartment. This time the reader, rather than Mrs. Dalloway, is invited to simulate the experience of that death in a highly embodied way. Such an embodied simulation may well be an unwelcome one, just as it is for the protagonist, who does not want to have that experience, nor does she want her guests to do so, either.

Embodied simulation hinges on the immediate and involuntary mirrored experience of, in this case, a fictional character’s horrific suicide. Such simulation also has, however, its “cognitive” aspects, insofar as it likely serves as the starting point for the reader’s extended reflections on the purpose of this violent and violating scene within Woolf’s narrative. Like Mrs. Dalloway, readers may feel that they are being subjected to an unwanted and possibly abusive narrative experience, and, if they continue to read, they may look to the narrative for explanations and/or justifications for its inclusion.

A few pages later, Woolf’s narrator offers additional clues to the possible purposes for that reported event and prompt to embodied simulation. The text reads:

32 Critics generally note Mrs. Dalloway’s identification with Septimus, rather than her lack of empathy. Elizabeth Abel writes, “It is a critical commonplace that Clarissa receives from Septimus a cathartic, vicarious experience of death that releases her to experience life’s pleasures more intensely” (1989, 38).

In a related vein, J. Hillis Miller has noted, “The death by suicide Woolf originally planned for Clarissa is fulfilled by Septimus, who dies for her, so to speak, a substitute suicide. Clarissa and Septimus seek the same thing: communication, wholeness, the oneness of reality, but only Septimus takes the sure way to reach it” (1982, 198).
She felt somehow very like him – the young man who had killed himself. She felt glad that he had done it; thrown it away. The clock was striking. The leaden circles dissolved in the air. He made her feel the beauty; made her feel the fun. But she must go back. She must assemble. She must find Sally and Peter. And she came in from the little room. (2005, 182)

Clarissa Dalloway’s resolution to rejoin her party raises more questions than it answers, yet it nevertheless points to the incongruous possibility that embodied simulation, of feeling the pain of another, may lead not only to identification (“she felt somehow very like him”), but also to some completely different psychic experience seemingly at odds with the initial simulation (“He made her feel the beauty; made her feel the fun.”) Bizarre, horrific, and/or luminous, this particular passage makes clear the potential variety of mirrored reactions, for they lead to or elide with other sensations, other emotions, other feelings. The same variety is likely to be apparent in the reader’s own reflections on this highly charged narrative experience. The experience of embodied simulation – Mrs. Dalloway’s, as well as the reader’s own – lies at the crux of the narrative. Woolf not only describes how it works, but also induces it in her readers.

*Dante and the Subjective Correlative*

As our second test case of liberated embodied simulation, we turn now to Dante’s famous sonnet “A ciascun’alma presa e gentil core,” which appears in the third section of his pre-*Commedia*, autobiographical prose-poem, the *Vita nuova*. There Dante the narrator recounts the origin of the first sonnet in the collection. After nine years, the poet sees his beloved Beatrice, who walks between two older women. As they pass each other in the street, their eyes meet, and she greets him:

e passando per una via, volse gli occhi verso quella parte ov’io era molto pauroso, e per la sua ineffabile cortesia, la quale è oggi meritata nel grande secolo, mi salutòe molto tanto virtuosamente, che me parve allora vedere tutti li termini della beatitudine. (1996, 15-16)

(and passing along a certain street, she turned her eyes to where I was standing faint-hearted and, with that indescribable graciousness for which today she is rewarded in the eternal life, she greeted me so miraculously that I seemed at that moment to behold the entire range of possible bliss.) (Musa 1973, 5)
Of note in this description are the narrator’s faint-heartedness (he is *pauroso*), the *
*ineffabile cortesia* (indescribable graciousness) with which the as-yet unnamed Beatrice greets him, and his blissful reaction to that entirely unexpected encounter (*me parve allora vedere tutti li termini de la beatitudine*, he writes). The poet leaves the scene as if intoxicated (*come inebriato*) and returns to his room. Falling asleep, he has a marvelous dream (*una maravigliosa visione*). The personified figure of Love appears to him in a flame-colored cloud; in his arms, the god carries the beloved, who is lightly veiled with a crimson cloth. In one hand, Love carries the burning heart of the poet and announces to the dreamer, “*Vide cor tuum*” (Behold thy heart). Then, waking Beatrice from her sleep, Love feeds her the poet’s heart: *e tanto si sforzava per suo ingegno, che le facea mangiare questa cosa che in mano li ardea, la quale ella mangiava dubitosamente* (19-20; and he forced her cunningly to eat of that burning object in his hand; she ate of it timidly (Musa 1973, 5).

The narrator then wakes up from his dream and decides to write a sonnet, not to Beatrice herself, but to an audience of *trovatori*, or love poets.³³ Sonnet I follows:

³³ “*Pensando io a·cciò che m’era apparuto, propuosi di farlo sentire a molti li quali erano famosi trovatori in quello tempo: e con ciò fosse cosa che io avessi già veduto per me medesimo l’arte del dire parole per rima, propuosi di fare uno sonetto, nel quale io salutasse tutti li fedeli d’Amore; e pregandoli che giudicassero la mia visione, scrissero a·l loro ciò che io avea nel mio sonno veduto. E cominciò allora questo sonetto*” (1996, 21-22; [Musa 1973, 6: Thinking about what I had seen, I decided to make it known to many of the famous poets of that time. Since just recently I had taught myself the art of writing poetry, I decided to compose a sonnet addressed to all of Love’s faithful subjects; and, requesting them to interpret my vision, I would write them what I had seen in my sleep]).
when suddenly Love appeared before me
[to remember how he really was appalls me].

Joyous, Love seemed to me, holding my heart
within his hand, and in his arms he had
my lady, loosely wrapped in folds, asleep.
He woke her then, and gently fed to her
the burning heart; she ate it, terrified.
And then I saw him disappear
in tears. (Musa 1973, 6)

There are many remarkable things about this poem, including (or especially) the range of embodied responses that the sonnet may evoke in the reader. Whatever they may be, those feelings are likely to register strongly. Dante chooses to explain his love by means of an imaginary scenario, a dream vision in which several impossible things occur: most notably, Love’s forceful yet humble feeding of the poet’s disembodied, flaming heart to Beatrice.

What do we make of this extraordinary image, which exemplifies what Giuseppe Mazzotta regards as the “radical metaphoricity” of the Vita nuova, a poem in which “substances are spiritualized, bodies are dematerialized, essences are incarnated in luminous bodies, physical surroundings have nothing opaque about them, ideas are refracted through images and sensuous qualities and lose their abstractions?” (1983, 10). Of note in Mazzotta’s characterization of the Vita nuova is the paradoxical relation of embodiment and disembodiment within Dante’s fantasies. We might reframe Mazzotta’s observation by noting the counter-factual nature of embodied experience, both in the representation itself and in the readers’ experiences of what Dante has presented to us. Love is given a body, and Dante’s flaming heart, removed from his body, is fed to a passive and receptive Beatrice, who eats that heart. Rather than asking what that sequence means, we ask what it feels like, or what feelings it might produce, and, too, whether an embodied narratology may shed light on our experiences of this poem.

Cognitive literary theorist Ellen Spolsky has explored the ways in which artists may represent the abstract, the unfamiliar, or the non-representable by a process of “transfiguration”: namely, engaging “the brain’s ability to move between the evidence of sense perception and introspected abstractions.”34 Through techniques of embodied representation, the artist enables readers (or viewers) to find partial matches between their own sensory experiences or memories on the one hand and, on the other, the abstract concepts that the artist also wishes to convey; the audience, in turn, produces representations within their own minds of things that might otherwise seem impossible to imagine.

In Sonnet I, we argue, following Spolsky, that Dante represents a somewhat abstract experience – passionate love from a distance enflamed by the gaze and the greeting of the beloved – as a paradoxically embodied one: Love’s feeding Beatrice the burning heart of

34 Spolsky offers a brilliant cognitive reading of Raphael’s painting The Transfiguration in order to theorize the ways in which the artist attempts to represent the unrepresentable – here, the mystery of Christ’s combined divinity and humanity. She writes, “The evolved human ability to move among structures of information, that is, knowledge, from different sources, and from the concrete to the abstract, is itself readable as a biology of transfiguration” (2007, 79).
Dante. As Albert Ascoli has astutely noted, Dante tells us that he is sending his poem out “for interpretation by a community of contemporary readers – specifically…the poets he calls the ‘fedeli d’amore,’ including Guido Cavalcanti, his ‘first friend.’” In doing so, Dante invites his readers to read him, as Ascoli explains, “both passively and actively.”

Ascoli quite rightly highlights what the critical tradition has largely neglected to study – i.e., that Dante asks his readers to interpret his dream, as the third line of the sonnet suggests (“in ciò che mi riscriva ’n suo parvente” [for your elucidation in reply]). Dante not only magnificently represents what is difficult to represent (love, or specifically, his love), but also requests that his audience tell him what he is feeling, and to validate his – the poet’s – feelings through their own perceptions. It is to the immediate, involuntary, and pre-rational experience of mirrored, empathic response that we now turn.

So what do we feel when we read the words “m’apparve Amor subitamente,/ cui essenza membrar mi dà orrore” (7-8). Do we feel horror, passion, or arousal? Do we giggle with delight, or cringe with dread? Of course, a number of responses are possible, as they are with the poem’s climactic sestet.

Arguably the sonnet turns on its final three lines, which present a hand-to-mouth gesture:

Poi la svegliava, e d’esto core ardendo
lei paventosa umilmente pascea.
Appresso gir lo ne vedea piangendo. (12-14; translated above)

The unfolding neuroscience of mirrored response does not tell us precisely what we are feeling in response to our own singular visualizations evoked by these lines, nor would most of us want it to. Neuroscience does, however, help us to understand why we might feel what we are feeling in response to this (or any other) poem, and also the simple presence of FoB alongside other, more “cognitive” responses that register in our awareness. The liberated embodied simulation evoked by the sonnet enables each reader, perhaps self-identifying as a fedele d’amore, first to feel and then to explore and elaborate a wide range of potential feelings – feelings that might, in fact, work against each other or that seem to be mutually exclusive.

For example, when Love feeds the beloved the heart umilmente, or “reverently,” how does that square with her reported experience of fear, or of his departure “in tears” (piangendo), which he conceals from her? What is this strange, imagined act, which seems both nurturing and over-bearing or forceful at the same time? Responses may vary according to the sex or normative gender of the reader (women may resonate more with the beloved than with Love, for example, and men with Love or the poet-witness, rather

---

35 Ascoli notes that Dante also broaches “lectio passio vel actus legendi” in the De vulgari eloquentia 2. 8. 3, and that “[t]he passage is particularly relevant to the present discussion of the interpenetrating roles of the author (singer) and reader.” Ascoli’s explanation offers a close parallel to the theory of intercorporeity (2008, 188, n. 33).

36 Guido Cavalcanti wrote a poem in response, and that was the beginning of the friendship between the two poets. However, no one guessed the meaning of the dream at the time, he says – namely, her impending death (Singleton 1949, 14).
than the beloved – though not necessarily). Readers may also sense the cross-gendered aspect of Love nursing Beatrice, like a reverse pietà. There is no single response, or set of responses evoked by the sonnet; rather, the poem produces, we suggest, an initial on-rush of involuntary feelings that the reader must then parse if he/she wishes to do so.

A Freudian reading of the poem would likely highlight the taboo and barely sublimated sexuality of the dream. The present reading, in contrast, foregrounds the shock of mirrored response to the bizarre and possibly moving scene of feeding that we attempt to visualize with our mind’s eye – a shock that may be categorized as erotic, appalling, delicious, disgusting, transcendent, or something else again – all together or separately.

In his 1922 essay “Hamlet and His Problems,” the poet T.S. Eliot developed a theory of emotion in art that seems oddly prescient:

The only way of expressing emotion in the form of art is by finding an ‘objective correlative;’ in other words, a set of objects, a situation, a chain of events which shall be the formula of that particular emotion; such that when the external facts, which must terminate in sensory experience, are given, the emotion is immediately evoked. (1969, 100)

Eliot sought to explain the writer’s predicament of representing that which is difficult to represent: i.e., emotion. A writer’s merely talking about emotion does not arouse it in the reader; instead, the writer must show, not tell, the reader what that emotion is by depicting “a set of objects, a situation, a chain of events” that “terminate in sensory experience” and evoke that emotion. Eliot’s concept of the “objective correlative” is essentially a method for triggering the phenomenon of embodied simulation. Artists and writers from Dante to Woolf and Eliot have long recognized this phenomenon, which they sought to activate within their readership through their own modes of artistic production.

Today, however, it might be more useful to reconceptualize Eliot’s signature technique as the “subjective correlative.” The objects, situation, or chain of events intended to convey and evoke emotion may be outside the characters in the text, but they

---

37 As Mark Musa comments, “It is difficult to understand the attitude of those critics who find sublimity in Love’s gruesome act of forcing the lady to eat the lover’s heart” (1973, 120).
38 Marianne Shapiro explains in her exploration of metonymy in Vita nuova that, “The dream, like the text, must be interpreted, or solved like a riddle. It is in Freudian terms the result of a transformation of ‘dream-thoughts’ (matter) into ‘dream content.’ This example seems to fall rhetorically within the tradition of courtly love and of the stil novo, but it is more than the record of oneiric domination by a woman. The Vita Nuova solemnizes (in virtue of being a dream) the great moments of her relation to the dreamer and launches into a new myth. The dream of the ‘eaten heart’ portends a new way of rhetorically embodying that myth” (1979, 109).
39 Indeed, these two examples, which hail from quite different time periods, cultures and genres, were selected precisely because they illustrate aesthetic uses of the Mirror Mechanism, which may be considered a transhistorical phenomenon. We would also suggest, however, that there must be a strongly historical and cultural component to liberated embodied simulation and to the artistic techniques used to encode and elicit intersubjective experiences; much more research needs to be done in order to historicize the phenomenon, however, and to integrate it with cultural studies.
need not be “objective” in any sense. Rather, they may be counter-factual and/or literally impossible (as Spolsky underscores with her concept of cognitive transfiguration); nevertheless they register within the subjectivity of the reader, as they are approximate correlates of our own experiences and embodied memories. Indeed, it was a high modernist fantasy (as well as a later New Critical one) to think that the literary work could be separated from the subjectivity of both author and reader, though that intimate relation could be defamiliarized through certain techniques of writing.

Dante’s goal in the *Vita nuova* was to write of Beatrice what had never been written of any other woman.\(^{40}\) He succeeded in doing so by inventing counter-factual embodied experiences standing for his unique love, which we explore through embodied cognition. It does not matter that each of our felt experiences may differ subtly or dramatically from those of others (a typical complaint against the “purely” subjective and hence irrelevant nature of readers’ responses\(^ {41}\)). There is never not reader response, at the core of which is the Feeling of Body. The author’s subjective correlative guides us into the imaginative lives of others, through the synesthesia of embodied metaphor – imagined bodily experience standing in for affective states and dispositions, for sensations, thoughts and memories, and liberated within the reading experience through the stillness of our own bodies. These things cannot be separated, though they can be creatively recombined, especially under the expert guidance of the writer.

**Conclusions**

In the present article we have proposed a new embodied approach to literary narratology, capitalizing upon recent neuroscientific findings on the role of the body in social cognition and language. Narrative is a peculiar form of mediated intersubjectivity, where the text – the medium – enables readers to entertain multiple relationships with fictional characters and through them with their author. Recently literary theorists have turned more and more to cognitive science, particularly drawing on the notion of Theory of Mind, and suggesting that many levels of textual creation, as well as analysis, can be explained by the peculiar meta-representational features of the human mind.

We believe that at best such cognitive literary theory constitutes only one of the many possible ways of relating to a literary text. Moreover, this approach neglects or minimizes what we take to be a fundamental element of both author’s and reader’s relation with the text: the Feeling-of-Body generated by liberated embodied simulation. Our proposal for an embodied narratology holds that such an approach can shed new light for literary studies by providing a parsimonious and unified account of both author’s and reader’s relation with the text.

The paradigm shift made possible by the new approach of embodied cognition provides a revisionary account of human uniqueness without divorcing it from nature, and thus moving beyond the untenable conflict between nature and nurture in a much more efficient way than with the two-step solution to the same problem proposed by

\(^{40}\) Noted in the last paragraph of the *Vita nuova*, and also discussed by Mazzotta (1983, 10).

\(^{41}\) For a more elaborate statement of this view, visit the *locus classicus*, Wimsatt and Beardsley’s “The Affective Fallacy” (1992).
classic cognitivism. According to classic cognitivism, all phenomena studied by the so-called *Geisteswissenschaften* should be first reduced to the mental events taking place within a (solipsistic) mind, eventually to be further investigated with the technological tools of neuroscience.

The program of naturalization endorsed by the embodied cognition approach starts from the assumption that human nature and its distinctive features, such as language, creativity, and art, can be fully grasped only when one adopts a non-solipsistic stance that emphasizes the social nature of human cognition. As the Italian philosopher Paolo Virno has recently argued,

[È] lecito postulare l’esistenza di un’intersoggettività preliminare, anteriore alla stessa formazione di soggetti distinti; la mente umana, a differenza di quanto suggerisce il solipsismo metodologico delle scienze cognitive, è originariamente pubblica o collettiva. (2010, 198)

(It is legitimate to postulate the existence of an originary intersubjectivity, which is anterior to formation of distinctive individual subjects. The human mind, contrary to what one might deduce from the methodological solipsism of the cognitive sciences, is originally public or collective.)

This approach, when combined with the evidence on the “we”-centric, pre-individual space instantiated by the shared resonance mechanisms we have reviewed above, can perhaps enable a naturalized version of humanism – a neurohumanism devoid of the negative connotations humanism stirred among scholars of the humanities in the second half of the last century, and based not on the Universal, but, as Virno proposes, on the Common.

Literary theory is emerging from a decades-long period of profound skepticism towards humanism, its disciplinary origins, and, more broadly, towards the hegemonic aspects of knowledge production presented in the name of humanism. That skepticism gave rise to a series of ideological critiques that few in the field would wish to set aside in an all-out rush toward a neo-humanistic neo-positivism. It is worth bearing in mind, however, that the sciences have their own longstanding tradition of skepticism, which is

---

42 Elaborating on the philosophical contributions of Duns Scotus and Gilbert Simondon, Virno argues against the supposed ontological primacy of the individual subject, proposing that what he refers to as “il Comune” (the Common) would ground both the emergence and development of human singular subjective individuality. This enables Virno to draw an interesting distinction between the “Universal,” defined as the way in which the human mind surreptitiously assigns a numerical entity to the Common, and the Common itself. Universal predicates, like “beautiful,” “intelligent,” “man,” etc., do not account for the *natura communis* that predates and makes individuation possible. The Common exists independently from its mental representation and at the same time makes it possible. The Common – for example, human nature – is not a predicate of individual human beings [like Hannah or Vittorio], but it nevertheless enables their own individuation as human subjects (206). Virno uses the logical and ontological heterogeneity between Common and Universal to draw a political distinction/alternative between Multitude and State.
manifested most strongly in a methodology based on the replicability and verifiability of experimental evidence. One can imagine (or re-imagine) a collaborative investigation of the human infused with both forms of skepticism, yet also with an openness to discovery, as literary studies and neuroscience explore the common ground that is embodied experience.

We would not wish to lose sight of our disciplinary differences during this exploration, nor could we easily do so. While scientists seek parsimonious answers to highly specific, testable questions, literary scholars tend, at least at this point in history, to search out a multiplicity of possible answers to questions often broader and more diffuse than those posed by scientists. While one discipline pulls toward polysemy, variety, and ambiguity, celebrating the inconclusive, the other seeks to resolve ambiguity and dispel uncertainty with verifiable answers to what are often real-world problems. Those very different attitudes might seem to put practitioners of these disciplines at cross-purposes with each other.

Yet they need not do so. As we move into a future that is changing more rapidly than we can imagine, it will be useful to draw on these and other modes of knowledge production and to hybridize them, as we have sought to do here. In this essay, we have attempted – from our different disciplinary vantage points – to bring our knowledges back to the body, and to remind ourselves, literally and literarily, just how we are feeling.

Acknowledgments

The authors wish to thank Pierpaolo Antonello, Albert R. Ascoli, Douglas Bruster, Patrick Colm Hogan, Renate Holub, Siri Hustvelt, Alfonso Iacono, Daniel Lochman, Carol Mackay, Marcella Rossman, Randolph Starn, the readers and staff of California Italian Studies and the graduate students of the Spring 2011 University of Texas course How Stories Make Us Feel. They also thank the English and Psychology Departments of UT for inviting V. G. to Austin. This work was supported by the EU grants ROSSI and TESIS to V. G.

Bibliography


Fish, S. 1980. *Is There a Text in This Class? The Authority of Interpretive Communities*. Cambridge, MA: Harvard University Press.


